

AlloyMax: Bringing Maximum Satisfaction to Relational Specifications

Changjian Zhang
Carnegie Mellon University
Pittsburgh, USA
changjiz@andrew.cmu.edu

Ryan Wagner
Carnegie Mellon University
Pittsburgh, USA
rrwagner@cs.cmu.edu

Pedro Orvalho
INESC-ID/IST, Universidade de Lisboa
Lisboa, Portugal
pedro.orvalho@tecnico.ulisboa.pt

David Garlan
Carnegie Mellon University
Pittsburgh, USA
garlan@cs.cmu.edu

Vasco Manquinho
INESC-ID/IST, Universidade de Lisboa
Lisboa, Portugal
vasco.manquinho@tecnico.ulisboa.pt

Ruben Martins
Carnegie Mellon University
Pittsburgh, USA
rubenm@andrew.cmu.edu

Eunsuk Kang
Carnegie Mellon University
Pittsburgh, USA
eskang@cmu.edu

ABSTRACT

Alloy is a declarative modeling language based on a first-order relational logic. Its constraint-based analysis has enabled a wide range of applications in software engineering, including configuration synthesis, bug finding, test-case generation, and security analysis. Certain types of analysis tasks in these domains involve finding an *optimal* solution. For example, in a network configuration problem, instead of finding any valid configuration, it may be desirable to find one that is most *permissive* (i.e., it permits a maximum number of packets). Due to its dependence on SAT, however, Alloy cannot be used to specify and analyze these types of problems.

We propose Alloy$^{\text{Max}}$, an extension of Alloy with a capability to express and analyze problems with optimal solutions. Alloy$^{\text{Max}}$ introduces (1) a small addition of language constructs that can be used to specify a wide range of problems that involve optimality and (2) a new analysis engine that leverages a *Maximum Satisfiability* (*MaxSAT*) solver to generate optimal solutions. To enable this new type of analysis, we show how a specification in a first-order relational logic can be translated into an input format of MaxSAT solvers—namely, a Boolean formula in *weighted conjunctive normal form* (*WCNF*). We demonstrate the applicability and scalability of Alloy$^{\text{Max}}$ on a benchmark of problems. To our knowledge, Alloy$^{\text{Max}}$ is the first approach to enable analysis with optimality in a relational modeling language, and we believe that Alloy$^{\text{Max}}$ has the potential to bring a wide range of new applications to Alloy.



This work is licensed under a Creative Commons Attribution International 4.0 License.

ESEC/FSE '21, August 23–28, 2021, Athens, Greece
© 2021 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-8562-6/21/08.
https://doi.org/10.1145/3468264.3468587

CCS CONCEPTS

• **Software and its engineering → Specification languages**; **Software system models**; **Formal methods**.

KEYWORDS

Alloy, SAT, MaxSAT, Relational specifications, Model synthesis

ACM Reference Format:
Changjian Zhang, Ryan Wagner, Pedro Orvalho, David Garlan, Vasco Manquinho, Ruben Martins, and Eunsuk Kang. 2021. AlloyMax: Bringing Maximum Satisfaction to Relational Specifications. In *Proceedings of the 29th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE '21), August 23–28, 2021, Athens, Greece.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3468264.3468587

1 INTRODUCTION

Alloy[16] is a declarative modeling language based on a first-order relational logic with transitive closure. Thanks to its expressive power, along with the analysis capability provided by its back-end engine, the Alloy Analyzer, it has been applied to a wide range of problems in software engineering, including protocol verification [4, 47], configuration analysis [27, 34], test-case generation [19], bug finding [9, 17], and security analysis [1, 18, 43].

The types of analysis that can be performed with a high-level modeling language like Alloy depend on the capability of its underlying engine. The Alloy Analyzer is at its core a *finite model finder*: Given a set of first-order logic (FOL) constraints that correspond to a system specification M and user query S (e.g., $S \equiv$ "find a valid network configuration"), the analyzer translates formula $M \wedge S$ to an equisatisfiable Boolean formula F such that a satisfying instance to F is also a solution to $M \wedge S$. This analysis is then performed by an off-the-shelf Boolean Satisfiability (SAT) solver. In general, there may be multiple satisfying instances F; the Alloy Analyzer also provides a way to *enumerate* these solutions by repeatedly invoking the solver with additional constraints that are intended to block the previously seen instances.

The reliance on the SAT-based analysis means, however, that the Alloy Analyzer cannot be used to automatically perform the following types of analysis:

- **Generation of optimal solutions**: It is sometimes desirable to find not only any solution but one that is *minimal* or *maximal*, depending on the problem being modeled. For instance, in the context of security exploit generation [43], it may be desirable to synthesize the *smallest* possible exploit that can demonstrate a security violation. In a network firewall configuration problem [34], a configuration setting that *maximizes* the number of allowed packets while satisfying a security policy is typically considered more desirable than a less permissive solution. However, since the Alloy Analyzer returns instances in an arbitrary order, finding such an optimal solution can only be done by enumerating the (typically very large) space of all instances.
- **Analysis with soft constraints**: Certain problems are naturally expressed using a combination of both *hard* and *soft* constraints, where a desirable solution should satisfy as many of the soft ones as possible. For example, a meeting scheduling problem may consist of hard constraints (e.g., "all participants must be present") as well as soft ones (e.g., "participants' time preferences"); if it is not possible to satisfy all of the latter, a solution that fulfills as many of these optional preferences is still more desirable than others. Since all of the constraints in Alloy are treated as hard constraints that must always be satisfied, it is currently not possible to analyze such types of problems.
- **Analysis with priorities**: When a problem has multiple, *incomparable* optimal solutions, it may be helpful to rank these solutions using some notion of *priority* among constraints. For instance, in the aforementioned scheduling problem, a participant may have a higher preference towards meeting times in the morning rather than in the afternoon. Again, since Alloy treats all of the constraints equally, it is not possible to solve problems where such a notion of priority plays an important role.

In this paper, we propose Alloy$^{\text{Max}}$, an extension of Alloy that overcomes these limitations by enabling an analysis of Alloy specifications with optimal solution generation. In particular, we propose both (1) a *language extension* to allow the user to specify soft constraints and indicate parts of an Alloy specification that must be optimized and (2) a *new analysis back-end* that leverages a *maximum satisfiability* (*MaxSAT*) solver to perform the analysis. To enable this analysis, we also provide a new translation mechanism from a high-level Alloy$^{\text{Max}}$ specification in FOL to an equivalent MaxSAT problem.

The proposed extension significantly extends the range of analyses that can be performed over Alloy specifications. It is also a strict generalization of Alloy: It introduces no semantic changes to existing Alloy specifications and includes only a small number of syntactic extensions, and thus should be easily adoptable by existing Alloy users.

We demonstrate the added analysis capability of Alloy$^{\text{Max}}$ on a variety of case studies, including exploit generation on microprocessors [43] (similar to the well-known Spectre [20] and Meltdown [24] attacks), graceful degradation in security [13], a wedding seating assignment [32], and a task scheduling problem [22]. We show that not only can our extension be used to perform analyses that were not possible in Alloy before, but the performance of our analysis engine is competitive to the existing Alloy Analyzer.

Table 1: The course timetable of the scheduling problem.

	Mon	**Tue**	**Wed**	**Thu**	**Fri**
AM	CS101	ML/Compiler	CS101	ML/Compiler	
PM	SE	OS	SE	OS	CS101

The contributions of this paper are as follows:

- An extension to Alloy, called Alloy$^{\text{Max}}$, and an accompanying analysis engine that can be used to generate optimal solutions and solve problems with soft, prioritized constraints (Section 2);
- A translation mechanism from a first-order relational logic to a weighted Boolean formula (Section 4);
- An optimization technique that uses high-level information in an input Alloy specification to guide a MaxSAT solver more efficiently towards optimal solutions (Section 5); and
- A collection of case studies demonstrating the new types of analyses in Alloy$^{\text{Max}}$, and a set of benchmark results demonstrating its performance on problems of varying sizes (Section 6).

2 EXAMPLES

As a motivating example, consider a model of a *course scheduling* problem in Alloy (Figure 1). A university offers a set of courses that students can register for each semester; in this model, we assume the school is offering five courses named CS101, Compiler, OS, ML, and SE (lines 6-10). Each lecture can take place during the morning (AM) or afternoon (PM) between Monday to Friday (lines 1-4). A set of all available lecture slots (lines 13-14) and their relations to the day and time are defined on lines 17-18, and the timetable for the courses (Table 1) is defined on line 20. Each student is associated with a set of *core* courses that are required for their major. For simplicity, we consider the scheduling problem for one particular student, Alice, who is assigned CS101 as a core course.

Each student's registration schedule must satisfy the following requirements:

- Each student must take at least three courses (line 35);
- Each student must take all of the *core* courses that they are assigned, and (line 36);
- A student cannot take courses whose lecture times conflict with each other (line 37).

Given this model, we can use a *run* command (line 47) in Alloy to generate a valid course schedule for Alice that satisfies all of the requirements. Figure 2a shows the first generated instance, indicating that Alice could take Compiler, CS101, OS, and SE to satisfy the requirements.

2.1 Solving Optimization Problems

Suppose that we wish to extend this model to encode students' interests in subjects (i.e., every student is interested in a set of courses). The goal is to be able to generate a schedule that *maximizes* the number of courses that match a student's interest, *in addition* to satisfying the three basic requirements. The following code snippet shows the new definition for Student and Alice, where Alice is interested in SE and ML.

```
1  abstract sig Day {}// The set of days for Mon to Fri.
2  one sig Mon, Tue, Wed, Thu, Fri extends Day {}
3  abstract sig Time {}// The set of times (AM and PM).
4  one sig AM, PM extends Time {}
5
6  abstract sig Course { // The set of courses.
7    // Each course has several lectures in a week.
8    lectures: set Lecture
9  }
10 one sig CS101, Compiler, OS, ML, SE extends Course {}
11
12 // In total, 5 (days) x 2 (AM/PM) lecture slots.
13 abstract sig Lecture { day: one Day, time: one Time }
14 one sig MonAM, MonPM, TueAM, ... extends Lecture {}
15 fact {
16   // The full day/time definition is omitted.
17   day = MonAM -> Mon + MonPM -> Mon + ...
18   time = MonAM -> AM + MonPM -> PM + ...
19   // The full timetable definition is omitted.
20   lectures = SE -> MonPM + SE -> WedPM + ...
21 }
22
23 abstract sig Student { // The set of students.
24   // Each student is assigned a set of core courses.
25   core: set Course,
26   // Each student may register for a set of courses.
27   courses: set Course
28 }
29 one sig Alice extends Student {} {
30   core = CS101 // Alice has to take CS101.
31 }
32 // The predicate for the three requirements.
33 pred validSchedule[courses: Student -> Course] {
34   all stu: Student {
35     #stu.courses > 2
36     stu.core in stu.courses
37     all disj c1,c2:stu.courses | not conflict[c1, c2]
38   }
39 }
40 pred conflict[c1, c2: Course] {
41   some l1, l2: Lecture {
42     l1 in c1.lectures and l2 in c2.lectures
43     l1.day = l2.day and l1.time = l2.time
44   }
45 }
46 // Generate a valid schedule.
47 run { validSchedule[courses] }
```

Figure 1: An Alloy model for the course scheduling system.

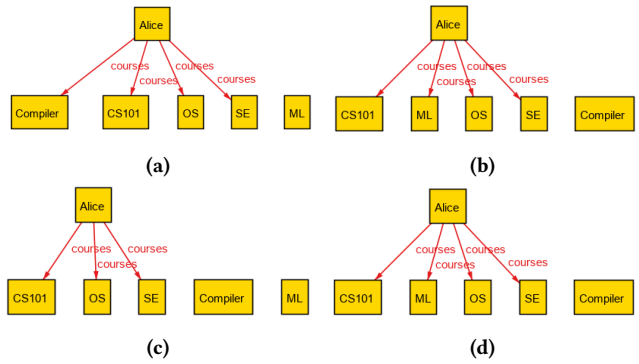


Figure 2: Instances of the course scheduling problem.

```
1  abstract sig Student {
2    core: set Course, courses: set Course,
3    // Courses that this student is interested in.
4    interests: set Course
5  }
6  one sig Alice extends Student {} {
7    core = CS101
8    interests = ML + SE // Interested in ML and SE.
9  }
```

Even though this maximization problem can be expressed in Alloy as follows, it contains a higher-order quantification (line 3, over possible relations from *Student* to *Course*), which cannot be solved by the Alloy Analyzer[1]:

```
1  run {
2    validSchedule[courses]
3    no courses': Student -> Course {
4      validSchedule[courses']
5      some stu:Student | #(stu.interests & stu.courses)
6        < #(stu.interests & stu.courses')
7    }
8  }
```

This optimization problem can be solved by our extension, Alloy$^{\text{Max}}$. The following code snippet shows a run command that uses our new *maxsome multiplicity* construct:

```
1  run MaxInterests1 {
2   validSchedule[courses]
3   all stu:Student| maxsome stu.interests & stu.courses
4  }
```

The meaning of the constraint on line 3 is that for any student, there exist *some* elements in the intersection of set *stu.interests* and set *stu.courses* that should be maximized, i.e., find an instance that maximizes the number of elements in this intersection.

An alternative way of specifying the optimization problem is by using our new *maxsome quantifier*:

```
1  run MaxInterests2 {
2    validSchedule[courses]
3    all stu: Student | maxsome comm: set Course |
4      comm in stu.interests and comm in stu.courses
5  }
```

Figure 2b shows the instance generated by a MaxSAT solver, which suggests that Alice could take CS101, ML, OS, and SE, where ML and SE are the courses that Alice is interested in.

2.2 Solving Problems with Soft Constraints

We wish to extend the model further to allow students to indicate preferences over lecture times for their courses. For example, suppose that Alice does not want to take courses on Thursday mornings and Friday afternoons. The following code snippet shows a run command with a constraint that encodes Alice's preferences. When the command is executed, the Alloy Analyzer fails to find any valid schedule (i.e., the resulting formula is unsatisfiable) because Alice has to take CS101, which has lectures on Friday afternoons.

```
1  run WithPrefer {
2   validSchedule[courses]
3   all stu:Student| maxsome stu.interests & stu.courses
4   // No lecture on Thu AM.
```

[1]Another extension of Alloy, called Alloy* [33], can be used to solve higher-order quantification, but does not support other types of analysis shown in Sections 2.2 and 2.3. We provide a more detailed comparison of Alloy$^{\text{Max}}$ and Alloy* in Section 7.

```
5   ThuAM !in Alice.courses.lectures
6   // No lecture on Fri PM.
7   FriPM !in Alice.courses.lectures
8  }
```

One way to find a satisfying instance is to relax some of the constraints; for instance, Alice could be asked to give up on some of her preferences to satisfy the requirements. However, Alloy does not provide a way to specify such desirable but *optional* constraints.

In Alloy$^{\text{Max}}$, we introduce a notion of *soft constraints*, i.e., a set of constraints that are not required but should be satisfied as much as possible. The following code shows a new type of construct called *soft fact*, which contains a list of soft constraints:

```
1  soft fact {
2    ThuAM !in Alice.courses.lectures
3    FriPM !in Alice.courses.lectures
4  }
5  run WithSoftPrefer {
6   validSchedule[courses]
7   all stu:Student| maxsome stu.interests & stu.courses
8  }
```

Given the above run command, Alloy$^{\text{Max}}$ will attempt to find an instance that maximizes the number of satisfied soft constraints, in addition to the other hard constraints that encode the three basic requirements. Figure 2c shows an instance where Alice could take CS101, OS, and SE, i.e., Alice gives up her Friday afternoons in order to register for CS101, but her Thursday mornings are freed up.

2.3 Solving with Priorities

However, Alice is not entirely satisfied with the above schedule. Although her Thursdays are free, she has to give up the ML course, which she is interested in. Given a choice between two conflicting goals—keeping Thursday mornings free and studying ML—she would rather choose the latter; i.e., ML has a higher *priority* for her. Currently, in Alloy, there is no way to express and analyze with such a notion of *priorities* among constraints.

In Alloy$^{\text{Max}}$, *priorities* can be explicitly assigned to maximal or soft constraints to indicate the user's preferences on which of the possibly conflicting constraints must be satisfied first. For example, in the following snippet, *maxsome[1]* indicates that maximizing the number of interested courses has a higher priority than satisfying the time preferences (i.e. *soft fact* is assigned default priority of 0).

```
1  soft fact {
2    ThuAM !in Alice.courses.lectures
3    FriPM !in Alice.courses.lectures
4  }
5  run WithSoftPreferAndPrior {
6    validSchedule[courses]
7    all stu: Student |
8      maxsome[1] stu.interests & stu.courses
9  }
```

Figure 2d shows the instance generated by the solver; i.e., Alice could take CS101, ML, OS, and SE, which maximizes her course interests over her time preferences.

In summary, our extension expands the range of analyses available in Alloy, including (1) generating an instance that maximizes or minimizes a given relation, which can be used for optimization problems (2) defining soft constraints and finding an instance that satisfies as many of these optional but desired constraints, and (3) specifying priorities among different soft or maximal constraints.

3 BACKGROUND ON MAXSAT

A propositional formula in *conjunctive normal form* (CNF) is defined as a conjunction of clauses, where each clause is a disjunction of literals such that a literal is either a propositional variable v_i or its negation $\neg v_i$. A clause is satisfied if at least one of its literals is satisfied. Finally, a CNF formula is satisfied if all its clauses are satisfied. Given a CNF formula ϕ, the Satisfiability (SAT) problem corresponds to deciding if there is an assignment such that ϕ is satisfied or proving that no such assignment exists.

The Maximum Satisfiability (MaxSAT) is an optimization version of the SAT problem. Given a CNF formula ϕ, the goal is to find an assignment that maximizes the number of satisfied clauses in ϕ. In partial MaxSAT, clauses in ϕ are split in hard ϕ_h and soft ϕ_s. Given a formula $\phi = (\phi_h, \phi_s)$, the goal is to find an assignment that satisfies all hard clauses in ϕ_h while minimizing the number of unsatisfied soft clauses in ϕ_s. This problem can be further generalized to weighted MaxSAT, where each soft clause has a positive weight, and the goal becomes to minimize the sum of the weights of unsatisfied soft clauses. MaxSAT algorithms have seen a remarkable improvement in the last decade [3] and can be used to solve problems in domains such as planning [48], data analysis [28], security [10], and bioinformatics [11]. In this paper, we leverage existing MaxSAT technology and extend Alloy to handle optimization problems.

4 SYNTAX AND SEMANTICS

This section describes the syntax and semantics of Alloy$^{\text{Max}}$. It also presents the formal translation rules from an Alloy$^{\text{Max}}$ specification to a MaxSAT problem.

4.1 Abstract Syntax

Figure 3 shows the abstract syntax of Alloy$^{\text{Max}}$. An Alloy$^{\text{Max}}$ problem is a tuple $P = \langle \mathcal{A}, D, F, F_S \rangle$ where $\mathcal{A}$ is the set of atoms in the universe, D is a set of relation declarations, F is an Alloy formula that defines the *hard* constraints that must be satisfied, F_S is a *soft* Alloy formula that defines the *soft* constraints that may or may not be satisfied. A (*hard*) Alloy formula may contain keywords such as *maxsome* for generating optimal solutions.

Optimization operators. Alloy$^{\text{Max}}$ extends the existing *some* multiplicity operator with *maxsome* and *minsome* for maximization and minimization, respectively. Informally, *some r* is true when relation *r* is not empty; but with *maxsome* (*minsome*), the solver returns a model with a non-empty *r* that has a maximal (minimal) number of tuples.

Alloy$^{\text{Max}}$ also extends the *no* multiplicity operator with *softno*. Informally, *no r* is true when *r* is empty. For *softno*, the solver tries to find a model with an empty *r*, but if no such satisfying instance exists, the solver instead returns a model where *r* has the minimal number of tuples. The difference between *minsome* and *softno* is that *minsome* requires the relation *r* to have at least one tuple.

In addition, the existential quantifier (i.e., *some*) is also extended with *maxsome* and *minsome*. In Alloy, *some* $e : r \mid F$ is true when there exists a tuple in relation *r* that makes formula *F* true. In Alloy$^{\text{Max}}$, when *maxsome* (*minsome*) is used in place of *some*, the

problem := univDecl relDecl* formula softFormula
softFormula := **soft** formula | **soft[0..]** formula | compositeSoft
compositeSoft := softFormula **and** softFormula

univDecl := {atom[,atom]*}
relDecl := rel :$_{\text{arity}}$
varDecl := var : expr
constant := {tuple*}
tuple := ⟨atom[,atom]*⟩

arity := 1 | 2 | 3 | ...
atom := identifier
rel := identifier

var := identifier

expr := rel | var | unary | binary | comprehension
unary := unop expr
unop := ~ | ^
binary := expr binop expr
binop := + | & | - | . | ->
comprehension := {varDecl | formula}

formula := elementary | composite | quantified
elementary := expr **in** expr | mult expr
mult := **some** | **maxsome** | **maxsome[0..]** | **minsome**
| **minsome[0..]** | **no** | **softno** | **softno[0..]** | **one**
composite := **not** formula | formula logop formula
logop := **and** | **or**
quantified := quantifier varDecl | formula
quantifier := **all** | **some** | **maxsome** | **maxsome[0..]**
| **minsome** | **minsome[0..]**

Figure 3: Abstract syntax of Alloy$^{\text{Max}}$. Bolded text are existing keywords in Alloy; text in blue are new keywords in Alloy$^{\text{Max}}$.

solver finds a model where r contains at least one tuple and tries to maximize (minimize) the number of tuples in r that make F true.

Soft constraints. Alloy$^{\text{Max}}$ introduces the *soft* fact keyword to specify a soft constraint, i.e., *soft fact* F where F is an Alloy formula but becomes *soft*. When multiple *soft fact* are present in a given specification, the overall soft constraint F_S is the conjunction of those soft Alloy formulas.

Priority. A priority $p \in \mathbb{N}_0$ may be specified along with every one of the new operators in Alloy$^{\text{Max}}$. When it is left unspecified, a default lowest priority 0 is assigned to the associated Alloy formula.

4.2 Translation

Overview. Figure 4 formally defines the translation rules for Alloy$^{\text{Max}}$ as an extension to the existing translation process in Alloy [15, 42]. Due to limited space, we focus on the parts of the translation process that are most relevant for Alloy$^{\text{Max}}$.

A relation in Alloy is translated into a matrix of Boolean variables (each of which is true if and only if the tuple represented by this particular variable is in the relation), and a relational expression (e.g., dot join) is represented by operations over one or more matrices. In Alloy, a FOL problem with bounds is translated into a Boolean formula, which is eventually converted into conjunctive normal form (CNF). In Alloy$^{\text{Max}}$, the idea is to instead convert it into *weighted* conjunctive normal form (WCNF), where (1) hard Boolean formulas are assigned the special weight of $+\infty$, to ensure that they are satisfied in every instance and (2) soft formulas are assigned different weights (depending on the user-specified priorities), to guide the solver towards an instance that maximizes the total sum of weights.

Translation Steps. First, an Alloy$^{\text{Max}}$ problem is translated into a Boolean formula with priorities, denoted by bool^p. A prioritized Boolean formula F^k is the Boolean formula F with an optional priority, k, which is used in a later part of the translation to assign weights to the corresponding WCNF clauses. When k is omitted, it represents a hard formula that must be satisfied.

In Figure 4, function M_S is used to translate soft constraints, i.e., constraints that are defined as part of *soft fact* in Alloy$^{\text{Max}}$. Given a soft constraint *soft fact* F with priority k, if F is a conjunction of Alloy formulas, M_S translates each conjuncts into a soft constraint also with priority k (line 17); otherwise, M_S translates the contained Alloy formula F by calling M and assigns k as its priority (line 18).

Function M is used to translate Alloy formulas. Keywords in vanilla Alloy follow the existing translation rules. For each of the new keywords in Alloy$^{\text{Max}}$ (e.g., *maxsome*, *minsome*), the translation rule is similar to the existing rule, except additional prioritized Boolean formulas are introduced to instruct the solver to search for a maximal or minimal solution. For instance, on line 21, *maxsome*$[k]$ p is translated by (1) stating that p must not be empty, as it is done for *some* in vanilla Alloy (line 20), and (2) including a prioritized Boolean formula for the presence of each tuple in p; given the corresponding WCNF, a MaxSAT solver attempts to find a solution that maximizes the number of tuples in p.

Each prioritized Boolean formula is transformed into a CNF formula with a priority for each clause by using the Tseitin transformation [44]. Finally, a CNF formula with priority is translated into a WCNF formula, to be solved by a MaxSAT solver. To guarantee that a soft clause with a higher priority is always satisfied first before the clauses with lower priorities, its weight is assigned to be greater than the sum of weights of all the clauses with smaller priorities. Formally, let C_i represent all the clauses with priority i. Then, their weight W_i is assigned as:

$$W_i = 1 + \sum_{j=0}^{i-1} W_j * |C_j| \tag{4.1}$$

where $|C_j|$ is the number of clauses of C_j and $W_0 = 1$. This kind of optimization is called *lexicographic optimization* and using this weight distribution is a known way of converting it to MaxSAT [29].

Example. We use an example to illustrate one of the translation rules (in particular, rule for *maxsome* on line 21 in Figure 4). Consider two relations $p : A \times B$ and $q : B \times C$, where $A = \{A1, A2\}$, $B = \{B1\}$, and $C = \{C1\}$. Suppose that we wish to find a model that

P : problem $\rightarrow$ bool^p (1)
R : relDecl $\rightarrow$ univDecl $\rightarrow$ matrix (2)
M : formula $\rightarrow$ env $\rightarrow$ bool^p (3)
M_S : softFormula $\rightarrow$ env $\rightarrow$ bool^p (4)
X : expr $\rightarrow$ env $\rightarrow$ matrix (5)
env: (quantVar $\cup$ relVar) $\rightarrow$ matrix (6)

$\overrightarrow{x}, \langle i_1, \ldots, i_k \rangle$: vectors (7)
$[m]$: matrix $\rightarrow$ $\{\langle int \rangle\}$, set of indices of matrix m (8)
$|m|$: matrix $\rightarrow$ dim, dimension of matrix m (9)
$\mathcal{M}$: dim $\rightarrow$ ($\langle$int$\rangle$ $\rightarrow$ bool^p) $\rightarrow$ matrix, constructor (10)
$\mathcal{M}(s^d, f) = \{m \mid |m| = s^d \wedge \forall \overrightarrow{x} \in \{0, \ldots, s-1\}^d, m[\overrightarrow{x}] = f(\overrightarrow{x})\}$ (11)
$\mathcal{M}$: dim $\rightarrow$ $\langle$int$\rangle$ $\rightarrow$ matrix, constructor (12)
$\mathcal{M}(s^d, \overrightarrow{x}) = \mathcal{M}(s^d, \lambda \overrightarrow{y}.$ if $\overrightarrow{y} = \overrightarrow{x}$ then $true$ else $false)$ (13)

$P[\mathcal{A}\ d_1, \ldots, d_n\ F\ F_S] =$ let $e = \bigcup_{i=1}^{n} (r_i \mapsto R[d_i]\mathcal{A})$ in $M[F]e \wedge M_S[F_S]e$ (14)
$R[r :_k]\mathcal{A} = \mathcal{M}(|\mathcal{A}|^k, \lambda \overrightarrow{x}$. freshVar()) (15)

$M_S[F_S$ and $G_S]e = M_S[F_S]e \wedge M_S[G_S]e$ (16)
$M_S[$soft[k] F and $G]e = M_S[$soft[k] $F]e \wedge M_S[$soft[k] $G]e$ (17)
$M_S[$soft[k] $F]e = (M[F]e)^k$, where F is not a conjunction (18)

$M[p$ in $q]e = \bigwedge(\neg X[p]e \vee X[q]e)$ (19)
$M[$some $p]e = \bigvee(X[p]e)$ (20)
$M[$maxsome[k] $p]e = M[$some $p]e \wedge \bigwedge(X[p]e)^k$ (21)
$M[$minsome[k] $p]e = M[$some $p]e \wedge \bigwedge(\neg X[p]e)^k$ (22)
$M[$one $p]e = ExactlyOne(X[p]e)$ (23)
$M[$no $p]e = \bigwedge(\neg X[p]e)$ (24)
$M[$softno[k] $p]e = \bigwedge(\neg X[p]e)^k$ (25)
$M[$not $F]e = \neg M[F]e$ (26)
$M[F$ and $G]e = M[F]e \wedge M[G]e$ (27)
$M[F$ or $G]e = M[F]e \vee M[G]e$ (28)
$M[$all $v : p \mid F]e =$ let $(m = X[p]e)$ in $\bigwedge_{\overrightarrow{x} \in [m]} (\neg m[\overrightarrow{x}] \vee M[F](e : v \mapsto \mathcal{M}(|m|, \overrightarrow{x})))$ (29)
$M[$some $v : p \mid F]e =$ let $(m = X[p]e)$ in $\bigvee_{\overrightarrow{x} \in [m]} (m[\overrightarrow{x}] \wedge M[F](e : v \mapsto \mathcal{M}(|m|, \overrightarrow{x})))$ (30)
$M[$maxsome[k] $v : p \mid F]e = M[$some $v : p \mid F]e \wedge \bigwedge(X[p]e)^k$ (31)
$M[$minsome[k] $v : p \mid F]e = M[$some $v : p \mid F]e \wedge \bigwedge(\neg X[p]e)^k$ (32)
$X[p + q]e = X[p]e \vee X[q]e$ (33)
$X[p \,\&\, q]e = X[p]e \wedge X[q]e$ (34)
$X[p - q]e = X[p]e \wedge \neg X[q]e$ (35)
$X[p \cdot q]e = X[p]e \cdot X[q]e$ (36)
$X[p \text{ -> } q]e = X[p]e \times X[q]e$ (37)
$X[\sim p]e = (X[p]e)^T$ (38)
$X[\hat{} p]e = IterativeSquare(X[p]e)$ (39)
$X[\{v : p \mid F\}]e =$ let $(m = X[p]e)$ in $\mathcal{M}(|m|, \lambda \overrightarrow{x}$. $m[x] \wedge M[F](e : v \mapsto \mathcal{M}(|m|, \overrightarrow{x})))$ (40)

Figure 4: Translation rules. bool^p is a Boolean formula with priority where $p \in \mathbb{N}_0$. When p is omitted, the formula represents a hard formula that must be satisfied. In other words, it has an infinite priority, i.e. $F = F^{+\infty}$.

satisfies the following formula:

$$maxsome\ p.q$$

Relation p is represented by a set of Boolean variables $\{p_{00}, p_{10}\}$ (where, for example, p_{00} is true if and only if $(A1, B1)$ is a tuple in p); similarly, q is represented by $\{q_{00}\}$. Then, the formula is translated into a prioritized Boolean formula as follows:

$$((p_{00} \wedge q_{00}) \vee (p_{10} \wedge q_{00})) \ \wedge (p_{00} \wedge q_{00})^0 \wedge (p_{10} \wedge q_{00})^0$$

where the first top-level conjunct says that $p.q$ must contain at least one tuple (equivalent to *some* $p.q$); the next two conjuncts are formulas with the lowest priority of 0 stating that $p.q$ should contain as many of the two possible tuples as possible.

To translate the above formula into CNF, the Tseitin transformation introduces fresh variables that represent sub-formulas:

$$a \leftrightarrow p_{00} \wedge q_{00}\ ,\ b \leftrightarrow p_{10} \wedge q_{00}$$

where $\{a, b\}$ can also be seen as the variables representing the tuples in relation $p.q$. Then, after the transformation, the following prioritized CNF is produced:

$$(a \vee b) \wedge a^0 \wedge b^0 \wedge$$
$$(\neg a \vee p_{00}) \wedge (\neg a \vee q_{00}) \wedge (a \vee \neg p_{00} \vee \neg q_{00}) \wedge$$
$$(\neg b \vee p_{10}) \wedge (\neg b \vee q_{00}) \wedge (b \vee \neg p_{10} \vee \neg q_{00})$$

Finally, we translate this formula into WCNF to be solved by a MaxSAT solver. Specifically, priorities are replaced with weights by Equation 4.1. Since a^0 and b^0 have priority 0, thus they are assigned weight 1 in the final WCNF formula.

4.3 Correctness of Translation

A *model* or an *instance* of an Alloy formula is an assignment of tuples to relations that makes the formula true. An Alloy problem may have multiple models. We call the set of all models the *solution space* of a problem. When two problems have the same solution space, we say that they are *semantically equivalent.*

Theorem 4.1. An Alloy$^{\text{Max}}$ problem $P = \langle \mathcal{A}, D, F, F_S \rangle$ is semantically equivalent to the problem $P' = \langle \mathcal{A}, D, F', true \rangle$ where F_S, the soft constraints, is replaced by $true$, and F' is the formula that results from replacing *maxsome* and *minsome* in F by *some*, and $softno$ by $true$.

Proof. To prove this theorem, we first state the following lemma:

Lemma 4.1. A prioritized Boolean formula F^k with a priority of a finite integer (i.e., a soft formula that can be satisfied or not) is semantically equivalent to $true$ in the corresponding non-prioritized Boolean formula.

According to this lemma, by replacing the soft formulas in the translation rules with $true$, we have:

- $soft[k]\ F$ (Figure 4, line 18) is semantically equivalent to $true$;

- $maxsome[k]\ p$ (line 21) and $minsome[k]\ p$ (line 22) are semantically equivalent to $some\ p$ (line 20);
- $softno[k]\ r$ (line 25) is semantically equivalent to $true$;
- $maxsome[k]\ v : p \mid F$ (line 31) and $minsome[k]\ v : p \mid F$ (line 32) are semantically equivalent to $some\ v : p \mid F$ (line 30).

Then, given an Alloy$^{\text{Max}}$ problem P, let P' be the problem that results from replacing all the soft facts with $true$, $maxsome$ and $minsome$ with $some$, and $softno$ formulas with $true$. Based on the above list of statements about semantic equivalence, it follows that a model is a solution to P if and only if it is also a solution to P'. □

Theorem 4.2. Given a MaxSAT solver that maximizes the total weight of a given WCNF formula, Alloy$^{\text{Max}}$ generates the optimal solution to the corresponding FOL problem.

Proof. This theorem can be proved by showing the correctness of each optimization operator. We show the correctness for $maxsome$ only; other operators can be proved similarly.

Given $maxsome[k]\ p$, where p is a relation, let $\{p_0, p_1, \ldots, p_n\}$ be the Boolean literals that represent the presence of tuples in p. Then, the formula will be translated into:

$$(p_0 \vee p_1 \vee \ldots \vee p_n) \wedge (p_0)^k \wedge (p_1)^k \wedge \ldots \wedge (p_n)^k$$

By maximizing the total weights of this formula, the number of satisfied clauses of form p_i (for $i \leq n$) is also maximized. Thus, Alloy$^{\text{Max}}$ returns a model where the size of the relation p is maximized.

□

5 OPTIMIZATION

By encoding problems in MaxSAT, Alloy$^{\text{Max}}$ can potentially take advantage of advances in MaxSAT technology to more efficiently solve the underlying optimization problem. However, when encoding the problem to a low-level WCNF, the high-level structure of the problem may be lost and become unavailable for MaxSAT algorithms to take advantage of. Prior approaches try to recover this information by clustering soft clauses into partitions and using them during solving [30, 37]. However, these approaches are heuristic-based and sometimes fail to capture meaningful relationships between soft clauses. In this paper, we propose to extract the partition information directly from the Alloy$^{\text{Max}}$ syntax and give this information to existing partition-based MaxSAT algorithms.

5.1 Partition-Based MaxSAT Solving

We briefly present an overview of unsatisfiability-based MaxSAT algorithms that use partition strategies [30, 37]. Unsatisfiability-based algorithms work by finding a sequence of unsatisfiable subformulas that correspond to increasing a lower bound on the number of falsified soft clauses until the formula becomes satisfiable and an optimal solution is found. These algorithms can be extended with partition strategies by taking as input the MaxSAT formula ϕ and partitioning this formula into formulas $\phi_1, \ldots, \phi_n$ with disjoint set of soft clauses. Different approaches have been proposed to partition a MaxSAT formula, but they are mostly based on having a graph representation of the formula and then using graph partitioning algorithms to create the partitions [30, 37].

The MaxSAT solver then takes as input these partitions and starts by solving ϕ_1. If the formula is unsatisfiable, then the lower bound on the number of falsified soft clauses is increased, and the formula is refined to allow one additional soft clause to be falsified. This loop is repeated until the formula becomes satisfiable. In this case, the algorithm found a solution to the formula that may or may not be optimal. If no more partitions are left, then this solution is guaranteed to be optimal. Otherwise, the next partition is added to the solver and this process is repeated until an optimal solution is found. Partitioning has the benefit of potentially finding smaller unsatisfiable subformulas at each iteration of the algorithm, which often leads to a sequence of unsatisfiable formulas that is easier to solve than without partitions. For further details on MaxSAT algorithms, we refer the interested reader to the literature [3].

5.2 Extracting Partitions from AlloyMax

Our approach, which we call *spec-based partitioning*, changes this existing process by using partitions that can be extracted directly from Alloy$^{\text{Max}}$ as the translation to MaxSAT is performed, rather than trying to guess "good" partitions at the WCNF level. In particular, partitions can be generated as follows.

- When adding a hard clause to the WCNF, it is assigned group 1. All the hard clauses in the WCNF belong to the same group.
- When translating an Alloy formula with optimization operators, i.e., *maxsome*, *minsome*, and *softno*, Alloy$^{\text{Max}}$ creates a new group g. Then, all the soft clauses associated with this optimization operator are assigned group g.
- When translating a soft constraint, i.e., *soft fact F*, Alloy$^{\text{Max}}$ creates a new group g' for it. Thus, each soft constraint is in its own group.

A common pattern that we leverage for creating partitions in Alloy$^{\text{Max}}$ is to use the *all* keyword. For example, consider set P and Q, and a relation $r : P \times Q$. For formula $all\ p : P \mid maxsome\ p.r$, the *all* ($\forall$) operator will be expanded to generate a $maxsome\ p.r$ formula for each $p \in P$. Each *maxsome* formula will be in a separate group, and we can easily partition the soft clauses by the tuples in set P. For example, recall the course scheduling problem in Section 2.2. The soft fact formula

```
1  soft fact {
2    ThuAM !in Alice.courses.lectures
3    FriPM !in Alice.courses.lectures
4  }
```

is first translated into two soft facts (Figure 4, line 17), and a group is created for each of these soft facts. In addition, the formula

```
1  all stu:Student| maxsome stu.interests & stu.courses
```

generates a group for each student. Then, the solver will first try to optimize each group individually and merge the results to find the global optimal solution.

In Section 6.3, we compare the graph-based partitioning strategy used in MaxSAT algorithms with Alloy$^{\text{Max}}$ partitioning and show the benefit of extracting partitions from a high-level model.

6 EVALUATION

We evaluate Alloy$^{\text{Max}}$ based on the following research questions:

RQ1: How useful is Alloy$^{\text{Max}}$ *for modeling different types of problems?* For usefulness, we focus on *applicability* and *scalability*.

RQ2: Is spec-based partitioning more efficient than no-partitioning and graph-based auto-partitioning?

RQ3: Is Alloy^{Max} *more efficient than Alloy* for generating optimal solutions?* Alloy* [33] is an extension that enables an analysis with higher-order quantifiers, so it could, in principle, be used to solve problems that generate an optimal solution as the one in Section 2.1. This question aims at comparing which approach is more efficient in this particular type of problem.

To answer these research questions, we apply Alloy^{Max} to five case studies from different domains. Section 6.2 gives brief introductions to each case study and presents the optimization goals.

6.1 Implementation

We implemented Alloy^{Max} based on Alloy 5.1.0, which, in turn, relies on Kodkod, a general-purpose relational model finder [42]. We extended Alloy to support the new syntax, and also modified Kodkod to support the new translation process. We use Open-WBO [31] as our backend MaxSAT with the same algorithm for both partitioning and non-partitioning. For spec-based partitioning, we extended the WCNF format to allow specifying group index at the beginning of each clause. We modified Open-WBO to read this new file format and changed the partition strategy to use the one specified in the file. All the experiments were run on a Linux machine with a 4 core 3.6GHZ CPU and 20 GB memory. Every problem was run 5 times (each with a 30 minutes timeout), from which the average was computed and reported in Table 2. The benchmark and tool are available for reproduction[2].

6.2 Case Studies

6.2.1 Course Scheduling. Recall the course scheduling model described in Section 2. In this benchmark, we randomly generate problems with C courses and S students; a course should have either *two* or *three* lectures among a week; each student should take at most R core courses and is interested in at most I courses.

The optimization goal is to maximize the number of interested courses that a student is registered for. The problem names in Table 2 follow the pattern $course_\{C\}_\{S\}_\{R\}_\{I\}$. The size of the Alloy^{Max} specifications ranges from 252 to 552 LOC.

6.2.2 CheckMate. CheckMate [43] is an automated tool based on Alloy for synthesizing proof-of-concept exploit code for hardware security attacks similar to Meltdown [24] and Spectre [20]. It uses a "micro- architecturally happens-before" graph [25], denoted as μhb, to model and analyze the execution process of micro-operations. In a CPU, an instruction goes through a sequence of stages to be executed (e.g., Fetch, Exec., and Commit). Then, a μhb graph represents the execution process of a series of program instructions.

An attack pattern (e.g., Meltdown or Spectre) is represented as a sub-graph pattern in μhb. CheckMate models the micro-architecture of a CPU, the events of instructions, and the attack pattern in Alloy, and uses the Alloy Analyzer to generate a μhb graph that contains the attack pattern. Moreover, CheckMate aims to synthesize security litmus tests, i.e., most *compact* programs that can demonstrate a security attack. However, since this optimization goal cannot be solved by Alloy, the authors of CheckMate developed an additional Python program to enumerate all the possible instances and find the litmus tests.

[2] https://github.com/SteveZhangBit/alloy-maxsat-benchmark

With Alloy^{Max}, this optimization task can be formulated as minimizing the relation that represents the μhb graph. In particular, this goal is modeled by using *softno*, *all* is used for leveraging partitioning (as shown below), and the returned instance contains the smallest μhb graph, which represents a security litmus test.

```
1  run {
2    ... // CheckMate constraints
3    all n: Node | softno n.uhb
4  }
```

The original CheckMate model is around 1,100 LOC in Alloy; to specify the optimization goal, it only involved adding the above *softno* formula to the specification.

6.2.3 Graceful Degradation. In network security, a simple way of mitigating an on-going attack is to shut down the entire system. In practice, however, this is unacceptable because the system will lose all of its functions. *Graceful degradation* [13, 46] is an approach for dynamically re-configuring parts of a system under an attack to maximize the remaining uncompromised functions. Typically, this is achieved by (1) shutting down only those components that have been compromised, and (2) replacing these components with backups (if available) to maintain the affected functions. This reconfiguration process also involves removing network connections to the compromised components and adding new ones to the backups.

Alloy can be used to model and analyze the problem of generating a valid network reconfiguration that maintains all system functions. However, depending on the extent of an attack, it may be impossible to restore some of those functions. To determine which of them can be restored, the Alloy user could manually comment out a selection of constraints that correspond to the functions, but the process to find an optimal solution (i.e., one that restores as many functions as possible) would be tedious.

With Alloy^{Max}, we use *soft fact* to model the functions as soft constraints (line 4-6). Given these, the solver will try to maximize the functions and give up on ones that are unsatisfiable. Moreover, in practice, every change to the architecture could have a cost, and the total cost (the number of architectural modifications) should also be minimized. Thus, we use *maxsome* to maximize the overlapping connections between the degraded architecture and the initial one (line 9), and use *softno* to minimize the new connections added to the degraded architecture (line 10). Furthermore, system function is considered more important than cost. Thus, soft constraints representing the functions are assigned a higher priority. The following snippet shows our optimization goals.

```
1  sig Component {
2    init: set Component, degraded: set Component
3  }
4  soft[1] fact {
5    NonCriticalFunction and CriticalFunction
6  }
7  run {
8    validArchitecture[degraded]
9    maxsome degraded & init
10   softno degraded - init
11 }
```

In our benchmark, each problem has C components including backup ones; the attacker's capability (in terms of the degree that the system might be compromised) is set to K, and; each problem has A initially compromised components. The problem names follow

the pattern *degrade_{C}_{K}_{A}*. The size of the specifications ranges from 240 to 326 LOC.

6.2.4 Wedding Table Seating. Consider a wedding table seating problem [32], where (1) each table seats a minimum and a maximum number of guests; (2) each person should be assigned to only one table; (3) each person is associated with a set of tags indicating their interests. The hard constraint is to generate a seating assignment for each person, and the optimization goal is to minimize the number of different tags between all people sitting at the same table.

Alloy can be used to model this problem and generate an arbitrary assignment. The following snippet shows relevant constraints.

```
sig Tag {}
sig Person { tags: set Tag }
sig Table { seat: set Person } {
  #seat =< MAX_NUM_PERSON// Maximum number of guests.
  #seat >= MIN_NUM_PERSON// Minimum number of guests.
}
// One seat for one person.
fact { all p: Person | one seat.p }
run {}
```

However, with this model, we cannot find the optimal solution that minimizes the number of different tags at a table. With Alloy$^{\text{Max}}$, this optimization goal can be modeled by using *softno*. In particular, to take advantage of the partitioning optimization, the goal is specified as, for each table, it should have no or minimized tags. The following expresses this optimization goal.

```
run TableBased { all t: Table | softno t.seat.tags }
```

For the benchmark, we randomly generated problems with G tags, T tables, and P people, where each table has minimum L and maximum H people, and each person has maximum M tags. The problem names follow the pattern *seat_{G}_{P}_{M}_{T}_{L}_{H}*. The size of the specifications ranges from 50 to 55 LOC.

6.2.5 Single Machine Scheduling (SMS). In many real-time systems, it is often impossible to schedule tasks to make them meet all their time requirements. In practice, one must schedule the tasks in such a way to can maximize the number of on-time tasks [22]. Consider a real-time system with a set of n tasks $\{\tau_1, \ldots, \tau_n\}$ to be executed on a uni-processor machine, i.e., only one task can be executed at a time. Each task is a triple $\langle r_i, p_i, d_i \rangle$ where r_i is the release time (i.e., the earliest start time), p_i the processing time, and d_i the deadline.

Tasks can be split into fragments so that one task can be interrupted in order to execute fragments of other tasks. However, the execution order of the fragments of one task should not be changed. Formally, a task τ_i has k_i fragments and each fragment f_i^j where $j \in \{1, \ldots, k_i\}$ has process time p_i^j, and we have $\sum_{j=1}^{k_i} p_i^j = p_i$. Finally, a task may depend on other tasks. Thus, if a task τ_i depends on task τ_j, τ_i must be executed only after τ_j has been completed.

Alloy can be used to model this problem and find an arbitrary subset of tasks that can complete within their deadlines. However, the maximality is not guaranteed. With Alloy$^{\text{Max}}$, we use *maxsome* to find the optimal schedule which maximizes the number of on-time tasks. The following snippet shows the optimization goal.

```
sig Task {} sig Completed in Task {}
run {
  CompleteOntime[Completed]
  maxsome Completed }
```

In our benchmark, we randomly generate problems with N tasks. The release time of each task is within a time frame R; each task has a maximum processing time P, and a maximum number of fragments F. A slack factor S is used to compute the deadline for each task where $d_i = r_i + random(1, S) * p_i$. Finally, a task depends on a maximum number of D tasks. The problem names follow the pattern *sms_{N}_{R}_{P}_{F}_{S}_{D}*. The size of the specifications ranges from 192 to 259 LOC.

6.3 Results

6.3.1 RQ1: How useful is Alloy$^{\text{Max}}$ *for modeling different types of problems?* We successfully modeled the aforementioned problems with Alloy$^{\text{Max}}$ (or migrate from existing ones). It only requires appending the optimization formulas to the original run commands in plain Alloy without any reconstruction.

Table 2, first column shows the problems used in our benchmark. The problem name reflects the scale of each problem. Specifically, the size of the Alloy$^{\text{Max}}$ specifications for these problems ranges from 50 LOC to 1,109 LOC. The number of variables in the generated WCNF ranges from 2,236 to 24,583,678, and the number of clauses ranges from 6,921 to 43,729,378. Although translating from a high-level language like Alloy$^{\text{Max}}$ may not produce the most succinct MaxSAT formulas, it can solve all these problems within a 30 minutes timeout.

6.3.2 RQ2: Is spec-based partitioning more efficient than no- partitioning and graph-based auto-partitioning? The results in Table 2 show that spec-based partitioning outperforms both no- and auto-partitioning in Course, CheckMate, and Seating problems. On SMS, spec-based partitioning shows little improvement over no-partitioning because these problems do not have a natural, problem-specific way of partitioning; as a result, the solver falls back to no-partitioning. On the other hand, spec-based partitioning has a negative impact on the first three Degradation problems. Moreover, we have the following observations and insights:

- When partitioning is applicable, spec-based partitioning outperforms auto-partitioning, because (1) the former does not suffer from the overhead introduced by the heuristic for guessing partitions, which can be significant for large problems, and the heuristic may fail and fall back to no-partitioning (e.g., in CheckMate and Degradation), and (2) auto-partitioning might fail to guess effective partitions (e.g., Seating problems where spec-based improves the performance but auto does not).
- When partitioning is not applicable (i.e., only one partition exists), spec-based partitioning falls back to no-partitioning. On the other hand, forcing auto-partitioning in this situation may have negative impact on performance (e.g., SMS).
- We hypothesize that partitioning works well if the optimal solution of a sub-problem (hard clauses plus some partitions) is part of the global optimal solution. The Course problems have this property, since there are no registration limits on the number of students for each course, and the registration for each student is completely independent of one another.
- We observed in the Degradation problems that if a partition added to the iterative process does not help refine the solution towards the optimal one, the overhead of the iteration may exceed the time of solving the problem at once. In other words,

Table 2: Results for comparing MaxSAT non-partitioning vs. MaxSAT partitioning vs. Alloy*. The numbers in the table are the solving time of the MaxSAT solver or the solver for Alloy* (i.e., excluding the translation time). There is no significant difference in translation time among these methods. * indicates auto-partition fails because the problem is too large and it falls back to no-partition. ** indicates only one partition exists in the problem and it falls back to no-partition.

Problem		No-Part. Solve (s)	Auto-Part. Solve (s)	Spec-Part. Solve (s)	Alloy* Solve (s)
Course	course_30_40_3_6	0.93	0.19	0.17	112.08
	course_40_50_3_6	3.90	0.42	0.31	437.72
	course_50_60_3_6	11.84	0.78	0.62	860.49
	course_60_70_3_6	31.38	1.14	0.90	1607.65
	course_70_80_3_6	74.83	1.89	1.47	T/O
	course_80_90_3_6	175.40	2.96	2.09	T/O
	course_90_100_3_6	277.93	4.72	3.09	T/O
CheckMate	Flush+Reload	123.01	133.60*	94.68	N/A
	Meltdown	353.74	368.13*	272.35	N/A
	Spectre	634.55	660.15*	520.91	N/A
Seating	seat_7_28_7_6_3_7	3.06	7.19	1.76	26.47
	seat_7_30_7_6_3_7	0.70	1.23	0.65	52.27
	seat_8_30_8_7_3_7	133.55	54.78	35.36	239.21
	seat_8_32_8_7_3_7	513.73	619.33	349.20	T/O
SMS	sms_34_20_6_3_3_2	20.06	31.00	20.05**	479.84
	sms_38_20_6_3_3_2	19.13	35.12	19.47**	863.60
	sms_42_20_6_3_3_2	21.71	52.12	20.63**	1464.60
	sms_46_20_6_3_3_2	186.98	263.31	184.93**	T/O
	sms_50_20_6_3_3_2	179.86	173.91	186.74**	T/O
	sms_52_20_6_3_3_2	81.22	198.73	85.21**	T/O
Degradation	degrade_10_2_1	0.26	0.77	0.28	N/A
	degrade_26_2_1	25.55	43.08*	27.76	N/A
	degrade_20_3_1	659.46	673.30*	T/O	N/A
	degrade_26_2_2	1442.41	1453.32*	898.7	N/A

partitioning works well when every sub-problem contributes towards the optimal solution.

6.3.3 RQ3: Is Alloy^{Max} *more efficient than Alloy* for generating optimal solutions?* As shown in Table 2, both with or without partitioning, Alloy^{Max} outperforms Alloy* on optimization problems that can be specified using higher-order quantifiers. Alloy* is a general-purpose higher-order solver (can be used to solve problems that Alloy^{Max} is not applicable to), and so its analysis method—based on counterexample-guided inductive synthesis (CEGIS) [40]—is not specifically targeted for solving optimization problems. On the other hand, since MaxSAT is designed for solving optimization, it is expected that Alloy^{Max} outperforms Alloy* on such problems.

We did not apply Alloy* to CheckMate because it would require significantly reconstructing the specification to express a higher-order formula that quantifies over all relevant relations to be searched (e.g., relations representing various properties of an attack graph to be minimized). Due to the complexity of this problem, there are 40 such relations, and so this would result in a formula with a quantifier depth of 40—not an Alloy expression that a user can reasonably be expected to write. In contrast, Alloy^{Max} introduces only one line of an optimization formula without changing the existing specification; this also shows the benefit of the built-in constructs in Alloy^{Max} for specifying optimization tasks.

Alloy* is not applicable to degradation problems, which contain soft constraints that cannot be specified in or solved by Alloy*.

6.3.4 Validation of Implementation. We manually checked the correctness of the results from Alloy^{Max} for the problems where the optimal solutions are known, including CheckMate, graceful degradation, and some subset of the course scheduling, seating, and SMS problems. Also, for those problems in Table 2 that can be solved by Alloy*, we checked that Alloy^{Max} and Alloy* produce the same optimal solutions.

6.3.5 Threats to Validity. Our selected case studies might not be representative enough to demonstrate the applicability of Alloy^{Max}. We believe that our benchmark contains a diverse set of examples from realistic domains (including exploit generation, network reconfiguration, and task scheduling) and of varying sizes (from relatively simple models like wedding seating to complex ones like CheckMate). However, we plan to continually expand our benchmark to explore other potential applications of Alloy^{Max} and further refine or modify the language as needed for new use cases.

In addition, the size of the problems in our benchmark might not be large enough to show how well Alloy^{Max} scales to other

Alloy problems in practice. Given that CheckMate is one of the most complex Alloy specifications available (based on our past experience with Alloy) and also has been used to generate practical attacks on real systems [43], we believe that our experiments show that Alloy^{Max} scales well to realistic problems. In addition, since its scalability depends on the efficiency of the underlying MaxSAT solver, Alloy^{Max} will likely benefit from further advances in MaxSAT technology, which is an active area of research [2].

7 RELATED WORK

Extensions to Alloy. Alloy* [33] is an Alloy extension that enables analysis of Alloy specifications with higher-order quantifiers. Although Alloy* is designed for a different purpose than Alloy^{Max}—mainly, enabling synthesis problems in Alloy—it could be used to encode certain types of optimization problems by using higher-order quantifiers, as shown in Section 2.1. However, as shown in Section 6.3, Alloy^{Max} outperforms Alloy* on these problems, in part because the former leverages MaxSAT solvers, which are designed for finding optimal solutions. In addition, Alloy^{Max} provides additional analyses involving soft constraints and priorities.

Aluminum [35] is an extension of Alloy that supports generation of *minimal instances*. Aluminum and Alloy^{Max} are designed for different purposes and differ on the notion of minimality. In particular, the former works by automatically removing as many tuples as possible such that the resulting instance remains a valid instance of the given Alloy specification, and allows the user to influence the order in which Alloy generates its next instances by augmenting the current instance with an additional tuple.

Cunha et. al [8] propose an approach called *target-oriented model finding*, where the goal is to allow the user to provide a partial specification of a *target* instance, and use a solver to generate an instance that is as *close* to the target as possible. Their work also leverages an extension of Kodkod with a MaxSAT solver as the underlying engine, although its overall goal is different from Alloy^{Max}.

Other types of solvers. There is a large body of work on *constrained optimization problems* (COPs)—a generalization of constraint satisfaction problems (CSPs) with an additional notion of an *objective function* to be optimized [38]. The state-of-the-art tools in this domain (e.g., Gurobi [12], CPLEX [14]) employ solving techniques that are different from MaxSAT, such as Branch and Cut algorithms [38]. Since the Alloy Analyzer already relies on SAT, we found MaxSAT to be a natural choice for bringing optimization to Alloy. Using a COP solver as an alternative backend to Alloy^{Max} is an interesting open problem, although it would involve a significantly different translation than the SAT-based one in Kodkod (which Alloy^{Max} builds on).

MiniZinc [36] is a constraint modeling language that is capable of solving various types of CSPs, including optimization. Built on top of solvers such Gurobi or CPLEX, MiniZinc is intended to be a high-level language for specifying CSPs and provides convenient built-in such as functions, data types, and global constraints. MiniZinc provides certain features (e.g., floating numbers) that are missing from Alloy^{Max}, while the relational core of Alloy is better suited for specifying complex structures that arise in software.

Answer set programming (ASP) [23] is a line of work on extending logic programming with capabilities for expressing and solving search problems, including those involving optimization [6]. Although both are declarative in nature, ASP and Alloy^{Max} support distinct styles of modeling: The logic programming paradigm underlying ASP is typically used for data and knowledge representation [7], while Alloy was designed for modeling software designs.

There are recent works on extending satisfiability modulo theories (SMT) solvers with optimization capabilities, such as OptiMathSAT [39], Symba [21], and Z3 [5]. These solvers could serve as an alternative backend for Alloy^{Max}. One benefit of this approach is that they operate over unbounded domains and for formulas involving decidable theories (e.g., integers), can provide a stronger analysis guarantee than a SAT-based backend. On the other hand, dedicated MaxSAT algorithms have shown to be more efficient than the SMT-based methods [26]. A more systematic investigation of these alternative backends for Alloy^{Max} remains future work.

8 CONCLUSION AND LIMITATIONS

We have presented Alloy^{Max}, an extension of the Alloy modeling language that leverages MaxSAT to solve optimization problems. We have demonstrated that despite introducing only a small number of new language constructs, Alloy^{Max} can be used to specify and solve a diverse set of problems that were previously not possible in Alloy. By enabling these new types of analyses, we believe that Alloy^{Max} has the potential to bring a wide range of new applications to Alloy, and its analysis capability will continue to grow as it benefits from rapid progress in the MaxSAT technology.

The translation rules shown in Section 4 support a richer semantics than the current syntax of Alloy^{Max} allows. For example, the rules support defining *soft* predicates (e.g., *soft* pred P[x: S]) and reuse them in other formulas. We plan to explore these richer semantics through further syntactic extensions (while also considering the possible risk of complicating the syntax) and investigate additional classes of problems that Alloy^{Max} could be applied to.

The current version of Alloy^{Max} supports only priorities among constraints. However, MaxSAT solvers generally use weights, which support more fine-grained optimizations. Adding weights to Alloy requires more subtle investigation on its semantic impact, but it could support a larger set of problems, such as modeling and maximizing utility functions (e.g., those used in self-adaptive systems planners [41] and automated data mismatch repair [45]).

ACKNOWLEDGMENTS

This work is supported in part by the National Science Foundation Awards CCF-1918140, CNS-1801546, and CCF-1762363, the Fundação para a Ciência e Tecnologia Awards PTDC/CCI-COM/31198/2-017, UIDB/50021/2020, and SFRH/BD/07724/2020, and the Office of Naval Research Award N00014172899. This material is based upon work supported by the NSA under Award No. H9823018D0008. This work is also supported by the Carnegie Mellon CyLab Security and Privacy Institute. Any views, opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the sponsoring agencies.

REFERENCES

[1] Devdatta Akhawe, Adam Barth, Peifung E. Lam, John Mitchell, and Dawn Song. 2010. Towards a Formal Foundation of Web Security. In *2010 23rd IEEE Computer Security Foundations Symposium*. 290–304. https://doi.org/10.1109/CSF.2010.27

[2] Fahiem Bacchus, Jeremias Berg, Matti Järvisalo, and Ruben Martins. 2020. *MaxSAT Evaluation 2020: Solver and Benchmark Descriptions*. Technical Report. University of Helsinki, Department of Computer Science.

[3] Fahiem Bacchus, Matti Järvisalo, and Ruben Martins. 2021. Maximum Satisfiability. In *Handbook of Satisfiability*, Armin Biere, Marijn Heule, Hans van Maaren, and Toby Walsh (Eds.). IOS Press, Chapter 24, 929 – 991. https://doi.org/10.3233/FAIA201008

[4] Hamid Bagheri, Eunsuk Kang, Sam Malek, and Daniel Jackson. 2015. Detection of Design Flaws in the Android Permission Protocol Through Bounded Verification. In *FM 2015: Formal Methods*, Nikolaj Bjørner and Frank de Boer (Eds.). Springer International Publishing, Cham, 73–89. https://doi.org/10.1007/978-3-319-19249-9_6

[5] Nikolaj Bjørner, Anh-Dung Phan, and Lars Fleckenstein. 2015. *νZ* - An Optimizing SMT Solver. In *Tools and Algorithms for the Construction and Analysis of Systems*, Christel Baier and Cesare Tinelli (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 194–199. https://doi.org/10.1007/978-3-662-46681-0_14

[6] Gerhard Brewka, Ilkka Niemelä, and Miroslaw Truszczynski. 2003. Answer Set Optimization. In *International Joint Conference on Artificial Intelligence (IJCAI)*, Georg Gottlob and Toby Walsh (Eds.).

[7] Stefano Ceri, Georg Gottlob, and Letizia Tanca. 1990. *Logic Programming and Databases*. Springer. https://www.worldcat.org/oclc/20595273

[8] Alcino Cunha, Nuno Macedo, and Tiago Guimarães. 2014. Target Oriented Relational Model Finding. In *Fundamental Approaches to Software Engineering*, Stefania Gnesi and Arend Rensink (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 17–31. https://doi.org/10.1007/978-3-642-54804-8_2

[9] Greg Dennis, Felix Sheng-Ho Chang, and Daniel Jackson. 2006. Modular Verification of Code with SAT. In *Proceedings of the 2006 International Symposium on Software Testing and Analysis* (Portland, Maine, USA) *(ISSTA '06)*. Association for Computing Machinery, New York, NY, USA, 109–120. https://doi.org/10.1145/1146238.1146251

[10] Yu Feng, Osbert Bastani, Ruben Martins, Isil Dillig, and Saswat Anand. 2017. Automated Synthesis of Semantic Malware Signatures using Maximum Satisfiability. In *Proceedings of the Annual Network and Distributed System Security Symposium*. The Internet Society. http://arxiv.org/abs/1608.06254

[11] Ana Graça, João Marques-Silva, and Inês Lynce. 2011. *Haplotype Inference Using Propositional Satisfiability*. Springer New York, New York, NY, 127–147. https://doi.org/10.1007/978-1-4419-6800-5_7

[12] Gurobi Optimization. 2021. Gurobi Optimizer Reference Manual. http://www.gurobi.com

[13] M P Herlihy and J M Wing. 1991. Specifying graceful degradation. *IEEE Transactions on Parallel and Distributed Systems* 2, 1 (1991), 93–104. https://doi.org/10.1109/71.80192

[14] International Business Machines Corporation (IBM). 2009. V12. 1: User's Manual for CPLEX.

[15] Daniel Jackson. 2000. Automating First-Order Relational Logic. In *Proceedings of the 8th ACM SIGSOFT International Symposium on Foundations of Software Engineering: Twenty-First Century Applications* (San Diego, California, USA) *(SIGSOFT '00/FSE-8)*. Association for Computing Machinery, New York, NY, USA, 130–139. https://doi.org/10.1145/355045.355063

[16] Daniel Jackson. 2006. *Software Abstractions: Logic, language, and analysis*. MIT Press.

[17] Daniel Jackson and Mandana Vaziri. 2000. Finding bugs with a constraint solver. In *International Symposium on Software Testing and Analysis (ISSTA)*. 14–25. https://doi.org/10.1145/347324.383378

[18] Eunsuk Kang, Aleksandar Milicevic, and Daniel Jackson. 2016. Multi-Representational Security Analysis. In *Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering* (Seattle, WA, USA) *(FSE 2016)*. Association for Computing Machinery, New York, NY, USA, 181–192. https://doi.org/10.1145/2950290.2950356

[19] Sarfraz Khurshid and Darko Marinov. 2004. TestEra: Specification-Based Testing of Java Programs Using SAT. *Autom. Softw. Eng.* 11, 4 (2004), 403–434. https://doi.org/10.1023/B:AUSE.0000038938.10589.b9

[20] Paul Kocher, Jann Horn, Anders Fogh, Daniel Genkin, Daniel Gruss, Werner Haas, Mike Hamburg, Moritz Lipp, Stefan Mangard, Thomas Prescher, Michael Schwarz, and Yuval Yarom. 2019. Spectre Attacks: Exploiting Speculative Execution. In *2019 IEEE Symposium on Security and Privacy (SP)*. 1–19. https://doi.org/10.1109/SP.2019.00002

[21] Yi Li, Aws Albarghouthi, Zachary Kincaid, Arie Gurfinkel, and Marsha Chechik. 2014. Symbolic Optimization with SMT Solvers. In *Proceedings of the 41st ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages* (San Diego, California, USA) *(POPL '14)*. Association for Computing Machinery, New York, NY, USA, 607–618. https://doi.org/10.1145/2535838.2535857

[22] Xiaojuan Liao, Hui Zhang, Miyuki Koshimura, Rong Huang, and Wenxin Yu. 2019. Maximum Satisfiability Formulation for Optimal Scheduling in Overloaded Real-Time Systems. In *PRICAI 2019: Trends in Artificial Intelligence*, Abhaya C. Nayak and Alok Sharma (Eds.). Springer International Publishing, Cham, 618–631. https://doi.org/10.1007/978-3-030-29908-8_49

[23] Vladimir Lifschitz. 2019. *Answer Set Programming*. Springer. https://doi.org/10.1007/978-3-030-24658-7

[24] Moritz Lipp, Michael Schwarz, Daniel Gruss, Thomas Prescher, Werner Haas, Stefan Mangard, Paul Kocher, Daniel Genkin, Yuval Yarom, and Mike Hamburg. 2018. Meltdown. arXiv:1801.01207 [cs.CR]

[25] Daniel Lustig, Michael Pellauer, and Margaret Martonosi. 2014. PipeCheck: Specifying and Verifying Microarchitectural Enforcement of Memory Consistency Models. In *2014 47th Annual IEEE/ACM International Symposium on Microarchitecture*. 635–646. https://doi.org/10.1109/MICRO.2014.38

[26] Inês Lynce, Vasco M. Manquinho, and Ruben Martins. 2018. Parallel Maximum Satisfiability. In *Handbook of Parallel Constraint Reasoning*, Youssef Hamadi and Lakhdar Sais (Eds.). Springer, 61–99. https://doi.org/10.1007/978-3-319-63516-3_3

[27] Ferney A. Maldonado-Lopez, Jaime Chavarriaga, and Yezid Donoso. 2014. Detecting Network Policy Conflicts Using Alloy. In *Abstract State Machines, Alloy, B, TLA, VDM, and Z*, Yamine Ait Ameur and Klaus-Dieter Schewe (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 314–317. https://doi.org/10.1007/978-3-662-43652-3_31

[28] Dmitry Malioutov and Kuldeep S. Meel. 2018. MLIC: A MaxSAT-Based Framework for Learning Interpretable Classification Rules. In *Principles and Practice of Constraint Programming*, John Hooker (Ed.). Springer International Publishing, Cham, 312–327. https://doi.org/10.1007/978-3-319-98334-9_21

[29] Joao Marques-Silva, Josep Argelich, Ana Graça, and Inês Lynce. 2011. Boolean lexicographic optimization: algorithms & applications. *Annals of Mathematics and Artificial Intelligence* 62, 3-4 (2011), 317–343. https://doi.org/10.1007/s10472-011-9233-2

[30] Ruben Martins, Vasco Manquinho, and Inês Lynce. 2013. Community-Based Partitioning for MaxSAT Solving. In *Theory and Applications of Satisfiability Testing – SAT 2013*, Matti Järvisalo and Allen Van Gelder (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 182–191. https://doi.org/10.1007/978-3-642-39071-5_14

[31] Ruben Martins, Vasco Manquinho, and Inês Lynce. 2014. Open-WBO: A Modular MaxSAT Solver. In *Theory and Applications of Satisfiability Testing – SAT 2014*, Carsten Sinz and Uwe Egly (Eds.). Springer International Publishing, Cham, 438–445. https://doi.org/10.1007/978-3-319-09284-3_33

[32] Ruben Martins and Justine Sherry. 2017. Lisbon Wedding: Seating arrangements using MaxSAT. *MaxSAT Evaluation 2017 : Solver and Benchmark Descriptions* B-2017-2 (2017), 25.

[33] Aleksandar Milicevic, Joseph P. Near, Eunsuk Kang, and Daniel Jackson. 2015. Alloy*: A General-Purpose Higher-Order Relational Constraint Solver. In *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*, Vol. 1. 609–619. https://doi.org/10.1109/ICSE.2015.77

[34] Sanjai Narain. 2005. Network Configuration Management via Model Finding. In *USENIX Conference on Systems Administration (LISA)*. 155–168.

[35] Tim Nelson, Salman Saghafi, Daniel J. Dougherty, Kathi Fisler, and Shriram Krishnamurthi. 2013. Aluminum: Principled scenario exploration through minimality. In *2013 35th International Conference on Software Engineering (ICSE)*. 232–241. https://doi.org/10.1109/ICSE.2013.6606569

[36] Nicholas Nethercote, Peter J. Stuckey, Ralph Becket, Sebastian Brand, Gregory J. Duck, and Guido Tack. 2007. MiniZinc: Towards a Standard CP Modelling Language. In *Principles and Practice of Constraint Programming – CP 2007*, Christian Bessière (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 529–543. https://doi.org/10.1007/978-3-540-74970-7_38

[37] Miguel Neves, Ruben Martins, Mikoláš Janota, Inês Lynce, and Vasco Manquinho. 2015. Exploiting Resolution-Based Representations for MaxSAT Solving. In *Theory and Applications of Satisfiability Testing – SAT 2015*, Marijn Heule and Sean Weaver (Eds.). Springer International Publishing, Cham, 272–286. https://doi.org/10.1007/978-3-319-24318-4_20

[38] Francesca Rossi, Peter van Beek, and Toby Walsh (Eds.). 2006. *Handbook of Constraint Programming*. Foundations of Artificial Intelligence, Vol. 2. Elsevier.

[39] Roberto Sebastiani and Patrick Trentin. 2020. OptiMathSAT: A Tool for Optimization Modulo Theories. *J. Autom. Reason.* 64, 3 (2020), 423–460. https://doi.org/10.1007/s10817-018-09508-6

[40] Armando Solar-Lezama, Liviu Tancau, Rastislav Bodik, Sanjit Seshia, and Vijay Saraswat. 2006. Combinatorial Sketching for Finite Programs. In *Proceedings of the 12th International Conference on Architectural Support for Programming Languages and Operating Systems* (San Jose, California, USA) *(ASPLOS XII)*. Association for Computing Machinery, New York, NY, USA, 404–415. https://doi.org/10.1145/1168857.1168907

[41] Clay Stevens and Hamid Bagheri. 2020. Reducing Run-Time Adaptation Space via Analysis of Possible Utility Bounds. In *2020 IEEE/ACM 42nd International Conference on Software Engineering (ICSE)*. 1522–1534.

[42] Emina Torlak and Daniel Jackson. 2007. Kodkod: A Relational Model Finder. In *Tools and Algorithms for the Construction and Analysis of Systems*, Orna Grumberg and Michael Huth (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 632–647.

https://doi.org/10.1007/978-3-540-71209-1_49

[43] Caroline Trippel, Daniel Lustig, and Margaret Martonosi. 2018. CheckMate: Automated synthesis of hardware exploits and security litmus tests. *Proceedings of the Annual International Symposium on Microarchitecture, MICRO* 2018-Octob (2018), 947–960. https://doi.org/10.1109/MICRO.2018.00081

[44] G. S. Tseitin. 1983. *On the Complexity of Derivation in Propositional Calculus*. Springer Berlin Heidelberg, Berlin, Heidelberg, 466–483. https://doi.org/10.1007/978-3-642-81955-1_28

[45] Perla Velasco-Elizondo, Vishal Dwivedi, David Garlan, Bradley Schmerl, and José Maria Fernandes. 2013. Resolving Data Mismatches in End-User Compositions. In *End-User Development*, Yvonne Dittrich, Margaret Burnett, Anders Mørch, and David Redmiles (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 120–136. https://doi.org/10.1007/978-3-642-38706-7_10

[46] Lingyu Wang, Sushil Jajodia, Anoop Singhal, Pengsu Cheng, and Steven Noel. 2013. k-Zero Day Safety: A Network Security Metric for Measuring the Risk of Unknown Vulnerabilities. *IEEE Transactions on Dependable and Secure Computing* 11, 1 (2013), 30–44. https://doi.org/10.1109/tdsc.2013.24

[47] Pamela Zave. 2017. Reasoning About Identifier Spaces: How to Make Chord Correct. *IEEE Transactions on Software Engineering* 43, 12 (2017), 1144–1156. https://doi.org/10.1109/TSE.2017.2655056

[48] Lei Zhang and Fahiem Bacchus. 2012. MAXSAT Heuristics for Cost Optimal Planning. *Proceedings of the AAAI Conference on Artificial Intelligence* 26, 1 (Jul. 2012). https://ojs.aaai.org/index.php/AAAI/article/view/8373