The Complexity of Constraints on Intervals and Lengths

Andrei Krokhin[1,⋆], Peter Jeavons[1,⋆], and Peter Jonsson[2,⋆⋆]

[1] Oxford University Computing Laboratory
Wolfson Building, Parks Road, Oxford OX1 3QD, UK
{Andrei.Krokhin,Peter.Jeavons}@comlab.ox.ac.uk
[2] Department of Computer and Information Science
Linköping University, S-581 83 Linköping, Sweden
Peter.Jonsson@ida.liu.se

Abstract. We study interval-valued constraint satisfaction problems (CSPs), in which the aim is to find an assignment of intervals to a given set of variables subject to constraints on the relative positions of intervals. One interesting question concerning such problems is to determine exactly how the complexity of an interval-valued CSP depends on the set of constraints allowed in instances. For the framework known as Allen's interval algebra this question was completely answered earlier by the authors.
Here we extend the qualitative framework of Allen's algebra with additional constraints on the lengths of intervals. We allow these length constraints to be expressed as Horn disjunctive linear relations, a well-known tractable and sufficiently expressive form of constraints. We completely characterize sets of qualitative relations for which the constraint satisfaction problem augmented with arbitrary length constraints of the above form is tractable. We also show that all the remaining cases are NP-complete.

1 Introduction and Summary of Results

A wide range of combinatorial search problems encountered in Computer Science and Artificial Intelligence can be naturally expressed as 'constraint satisfaction problems' [1], in which the aim is to find an assignment of *values* to a given set of *variables* subject to specified *constraints*. For example, the standard propositional SATISFIABILITY problem [2] may be viewed as a constraint satisfaction problem where the variables must be assigned Boolean values, and the constraints are specified by clauses. Further examples include GRAPH COLORABILITY, CLIQUE, and BANDWIDTH problems, scheduling problems, and many others (see [3, 4]).

⋆ Partially supported by the UK EPSRC grant GR/R29598.
⋆⋆ Partially supported by the Swedish Research Council (VR) under grant 221-2000-361.

Constraints are usually specified by means of relations. Hence the general constraint satisfaction problem can be parameterised according to the relations allowed in an instance. For any set of relations $\mathcal{F}$, the class of constraint satisfaction problem instances where the constraint relations are all members of $\mathcal{F}$ is denoted $CSP(\mathcal{F})$. The most well-known examples of such parameterised problems are Generalized Satisfiability [5], where the parameter is the set of allowed Boolean relations, and Graph H-coloring [6], where the parameter is the single graph H.

In studying CSPs over infinite sets of values, arguably the most important type of problem is when the constraints are specified by binary relations and the set of possible values for the variables is the set of intervals on the real line. Such problems arise, for example, in many forms of temporal reasoning [7, 8, 9, 10], where an event is identified with the interval during which it occurs. They also arise in computational biology, where various problems connected with physical mapping of DNA lead to interval-valued constraints [11, 12, 13]. Interval-valued CSPs can be naturally augmented with constraints on the lengths of the intervals, and the complexity of such extended problems will be the object of our main interest in this paper.

Before we describe our new results, we first discuss four closely related families of problems involving intervals which have previously been studied.

The prototypical problem from the first family is the Interval Graph Recognition problem [14]. An *interval graph* is an undirected graph such that there is assignment of intervals to the nodes with two nodes adjacent if and only if the two corresponding intervals intersect. Given an arbitrary graph G, the question of deciding whether G is an interval graph is rarely viewed as a constraint satisfaction problem, but in fact it is easily formulated as such a problem in the following way: every pair of adjacent nodes is constrained by the relation $r =$"intersect" over pairs of intervals, and every pair of non-adjacent nodes is constrained by the complementary relation $\bar{r} =$"disjoint". This fundamental Interval Graph Recognition problem is tractable, and it also remains tractable if we impose additional constraints on the lengths of the intervals which require all intervals to be of the same length (the Unit Interval Graph Recognition problem [15]). In contrast, it was shown in [16] that if we allow boundaries to be specified for the lengths of intervals, or even exact lengths (which are not necessarily all equal), then the corresponding problems (called Bounded Interval Graph Recognition and Measured Interval Graph Recognition, respectively) are NP-complete.

A typical problem from the second family is the Interval Graph Sandwich problem [12]. Given two graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ such that $E_1 \subseteq E_2$, the question is whether there is an interval graph $G = (V, E)$ with $E_1 \subseteq E \subseteq E_2$. Clearly, this is a generalization of the corresponding recognition problem (the case when $E_1 = E_2$). The Interval Graph Sandwich problem can be represented as a constraint satisfaction problem as follows: to any $e \in E_1$ assign the constraint $r =$"intersect", to any $e \notin E_2$ assign the constraint $\bar{r} =$"disjoint", and leave all pairs of variables corresponding to edges from $E_2 \setminus E_1$

unrelated. This problem was shown to be NP-complete along with the UNIT INTERVAL GRAPH SANDWICH problem where all intervals are required to be of the same length [12].

Other examples of problems from the first two families include the CIRCLE GRAPH RECOGNITION problem [17] and the CIRCLE GRAPH SANDWICH problem [18]. These problems can be formulated as a constraint satisfaction problem in the same way as above using the constraint relation $r =$"overlap".

The third family of problems we mention is the so-called INTERVAL SATISFIABILITY problems [8, 19, 20]. In these problems every pair of interval variables is again constrained in some way, but the constraints this time are chosen from a given set $\mathcal{F}$ of relations. In [8, 19, 20] only a small number of possibilities for $\mathcal{F}$ are considered. It is shown there that for some choices of $\mathcal{F}$ the resulting problem is tractable, whilst for others it is NP-complete. The complexity of INTERVAL SATISFIABILITY with all intervals of the same length is also studied in [19].

The fourth type of problem we mention is the satisfiability problem for *Allen's interval algebra* [7], denoted $\mathcal{A}$-SAT. Allen's algebra contains 13 basic relations (corresponding to the 13 ways two given intervals can be related from a "qualitative" point of view). The set $\mathcal{A}$ contains not just these basic relations, but all $2^{13} = 8192$ possible unions of them. The problems $\mathcal{A}$-SAT($\mathcal{F}$) are similar to problems of the third type above, except that unrelated pairs of variables are allowed. They can also be represented as INTERVAL SATISFIABILITY with $\mathcal{F}$ being an arbitrary subset of $\mathcal{A}$ containing the total relation. The complexity of problems of the form $\mathcal{A}$-SAT($\mathcal{F}$) has been intensively studied in the Artificial Intelligence community (see, e.g., [10, 21, 22]), and a complete classification of the complexity of such problems was obtained in [9]. In that paper it is shown that there are exactly 18 maximal tractable fragments of $\mathcal{A}$, and for any subset $\mathcal{F}$ not entirely contained in one of those the problem $\mathcal{A}$-SAT($\mathcal{F}$) is NP-complete.

Many problems, e.g., certain scheduling problems, can conveniently be expressed as $\mathcal{A}$-SAT($\mathcal{F}$) with additional constraints on the lengths of the intervals. Moreover, in [3], it was suggested that many important forms of constraints on lengths can be expressed in the form of Horn disjunctive linear relations. This class of relations is known to be tractable [23, 24] and at the same time it allows us to express all elementary constraints such as fixing the length, or bounding the length of an interval by a given number, or comparing the lengths of two intervals. It was proved in [3] that only three out of the 18 maximal tractable fragments for $\mathcal{A}$-SAT($\mathcal{F}$) preserve tractability when extended with Horn disjunctive linear constraints on lengths; the other 15 become NP-complete. In this paper we study how we need to further restrict those 15 fragments to obtain tractable cases. The main result is a complete classification of complexity for $\mathcal{A}$-SAT($\mathcal{F}$) with additional constraints on lengths. We show that such problems are either tractable or strongly NP-complete. Moreover, we give a complete description of the tractable cases, which allows one to easily determine whether a given set $\mathcal{F}$ falls into one of the tractable cases.

As well as giving a complete classification, our result also establishes a new dichotomy theorem for complexity. Dichotomy theorems are results concerning a class of related problems (with some parameter) which assert that, for some values of the parameter, the problems in the class are tractable while for all other values they are NP-complete. Such theorems are of interest because it is well known [2] that if P≠NP then, within NP, there are infinitely many pairwise inequivalent problems of intermediate complexity. Dichotomy results rule out such a possibility within certain classes of problems.

Dichotomy theorems have previously been established for the GENERALIZED SATISFIABILITY [5] and GRAPH H-COLORING [6] problems mentioned above as well as the DIRECTED SUBGRAPH HOMEOMORPHISM problem [25].

Constraint satisfaction problems have been a fruitful source of dichotomy results (see, e.g., [26, 27]). For constraint satisfaction problems, the relevant parameter is usually the set of relations, $\mathcal{F}$, specifying the allowed constraints. This parameter usually runs over an infinite set of values. In the case of Allen's algebra, even though the number of different values for $\mathcal{F}$ is finite, it is astronomical ($2^{8192} \approx 10^{2466}$) which excludes the possibility of computer-aided exhaustive case analysis.

The usual tool for proving dichotomy theorems is reducibilty via expressibility. This is done by showing that one set of relations expresses another, so that one problem can be reduced to the other. This is the method used in [5, 9, 26], and a similar method is used here. After identifying certain tractable fragments, we find some NP-complete fragments and then show how any subset not entirely contained in one of the tractable sets can express some already known NP-complete fragment.

The full version of the present paper is available in electronic form [28].

2 Preliminaries and Background

Allen's interval algebra [7], denoted $\mathcal{A}$, is a formalism for expressing *qualitative* binary relations between intervals on the real line. By "qualitative" we mean "invariant under all continuous injective monotone transformations of the real line". An interval x is represented as a pair $[x^-, x^+]$ of real numbers with $x^- < x^+$, denoting the left and right endpoints of the interval, respectively. The qualitative relations between intervals are the $2^{13} = 8192$ possible unions of the 13 *basic interval relations*, which are shown in Table 1. For the sake of brevity, relations between intervals will be written as collections of basic relations, omitting the sign of union. So, for instance, we write (pmf^{-1}) instead of $\mathsf{p} \cup \mathsf{m} \cup \mathsf{f}^{-1}$.

The problem of *satisfiability* ($\mathcal{A}$-SAT) in Allen's algebra is defined as follows.

Definition 1. *Let $\mathcal{F} \subseteq \mathcal{A}$ be a set of interval relations. An instance I of $\mathcal{A}$-SAT($\mathcal{F}$) over a set, V, of variables is a set of constraints of the form xry where $x, y \in V$ and $r \in \mathcal{F}$. The question is whether I is* satisfiable*, i.e., whether there exists a function, f, from V to the set of all intervals such that $f(x)\, r\, f(y)$ holds for every constraint xry in I. Any such function f is called a* model *of I.*

Allen's interval algebra $\mathcal{A}$ consists of the 8192 possible relations between intervals together with three standard operations on binary relations: *converse*

Table 1. The thirteen basic relations. The endpoint relations $x^- < x^+$ and $y^- < y^+$ that are valid for all relations have been omitted.

Basic relation		Example	Endpoints
x precedes y	p	xxx	$x^+ < y^-$
y preceded by x	p^{-1}	yyy	
x meets y	m	xxxx	$x^+ = y^-$
y met by x	m^{-1}	yyyy	
x overlaps y	o	xxxx	$x^- < y^- < x^+$,
y overlapped by x	o^{-1}	yyyy	$x^+ < y^+$
x during y	d	xxx	$x^- > y^-$,
y includes x	d^{-1}	yyyyyyy	$x^+ < y^+$
x starts y	s	xxx	$x^- = y^-$,
y started by x	s^{-1}	yyyyyyy	$x^+ < y^+$
x finishes y	f	xxx	$x^+ = y^+$,
y finished by x	f^{-1}	yyyyyyy	$x^- > y^-$
x equals y	$\equiv$	xxxx yyyy	$x^- = y^-$, $x^+ = y^+$

$\cdot^{-1}$, *intersection* $\cap$ *and composition* $\circ$. Subsets of $\mathcal{A}$ that are closed under the operations of converse, intersection, and composition are said to be *subalgebras*. For a given subset $\mathcal{F}$ of $\mathcal{A}$, the smallest subalgebra containing $\mathcal{F}$ is called the subalgebra *generated* by $\mathcal{F}$ and is denoted by $\langle\mathcal{F}\rangle$. It is easy to see that $\langle\mathcal{F}\rangle$ is obtained from $\mathcal{F}$ by adding all relations that can be obtained from the relations in $\mathcal{F}$ by using the three operations of $\mathcal{A}$.

It is known [10], and easy to prove, that, for every $\mathcal{F} \subseteq \mathcal{A}$, the problem $\mathcal{A}$-SAT$(\langle\mathcal{F}\rangle)$ is polynomially equivalent to $\mathcal{A}$-SAT$(\mathcal{F})$. Therefore, to classify the complexity of $\mathcal{A}$-SAT$(\mathcal{F})$ it is only necessary to consider these problems for *subalgebras* of $\mathcal{A}$. Throughout the paper, $\mathcal{S}$ denotes a subalgebra of $\mathcal{A}$.

In the following we shall use the symbol $\pm$, which should be interpreted as follows. A condition involving $\pm$ means the conjunction of two conditions: one corresponding to $+$ and one corresponding to $-$. For example, condition $r \cap (\mathsf{dsf})^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{d})^{\pm 1} \subseteq r$ means that both $r \cap (\mathsf{dsf}) \neq \emptyset \Rightarrow (\mathsf{d}) \subseteq r$ and $r \cap (\mathsf{d}^{-1}\mathsf{s}^{-1}\mathsf{f}^{-1}) \neq \emptyset \Rightarrow (\mathsf{d}^{-1}) \subseteq r$ hold. The main advantage of using the $\pm$ symbol is conciseness: in any subalgebra of $\mathcal{A}$, the '+' and the '$-$' conditions are satisfied (or not satisfied) simultaneously, and, therefore only one of them needs to be verified.

The following notation is used in Fig. 1 to describe some important subalgebras of $\mathcal{A}$. Let $q_0 = (\mathsf{pmo})$. Further, let $q_1 = q_0 \cup (\mathsf{d}^{-1}\mathsf{f}^{-1})$, $q_2 = q_0 \cup (\mathsf{df})$, $q_3 = q_0 \cup (\mathsf{df}^{-1})$, $q_4 = q_0 \cup (\mathsf{ds})$, $q_5 = q_0 \cup (\mathsf{d}^{-1}\mathsf{s}^{-1})$, $q_6 = q_0 \cup (\mathsf{d}^{-1}\mathsf{s})$.

A complete classification of complexity of problems of the form $\mathcal{A}$-SAT$(\mathcal{F})$ was obtained in [9].

Theorem 1 **([9]).** *For any subset $\mathcal{F}$ of $\mathcal{A}$, either $\mathcal{A}$-SAT$(\mathcal{F})$ is NP-complete or $\mathcal{F}$ is included in $\mathcal{S}$, where $\mathcal{S}$ is one of $\mathcal{S}_{\mathsf{p}}$, $\mathcal{S}_{\mathsf{d}}$, $\mathcal{S}_{\mathsf{o}}$, $\mathcal{E}_{\mathsf{p}}$, $\mathcal{E}_{\mathsf{d}}$, $\mathcal{E}_{\mathsf{o}}$, $\mathcal{A}_i (1 \le i \le 4)$, $\mathcal{B}_i (1 \le i \le 4)$, $\mathcal{S}^*$, $\mathcal{E}^*$, $\mathcal{A}_{\equiv}$, $\mathcal{H}$ (see Fig. 1), for which $\mathcal{A}$-SAT$(\mathcal{S})$ is tractable.*

$$\begin{array}{ll}
\mathcal{S}_{\mathsf{p}} = \{r \mid r \cap q_1^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{p})^{\pm 1} \subseteq r\} & \mathcal{E}_{\mathsf{p}} = \{r \mid r \cap q_4^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{p})^{\pm 1} \subseteq r\} \\
\mathcal{S}_{\mathsf{d}} = \{r \mid r \cap q_1^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{d}^{-1})^{\pm 1} \subseteq r\} & \mathcal{E}_{\mathsf{d}} = \{r \mid r \cap q_4^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{d})^{\pm 1} \subseteq r\} \\
\mathcal{S}_{\mathsf{o}} = \{r \mid r \cap q_1^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{o})^{\pm 1} \subseteq r\} & \mathcal{E}_{\mathsf{o}} = \{r \mid r \cap q_4^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{o})^{\pm 1} \subseteq r\} \\
\mathcal{A}_1 = \{r \mid r \cap q_1^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{s}^{-1})^{\pm 1} \subseteq r\} & \mathcal{B}_1 = \{r \mid r \cap q_4^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{f}^{-1})^{\pm 1} \subseteq r\} \\
\mathcal{A}_2 = \{r \mid r \cap q_1^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{s})^{\pm 1} \subseteq r\} & \mathcal{B}_2 = \{r \mid r \cap q_4^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{f})^{\pm 1} \subseteq r\} \\
\mathcal{A}_3 = \{r \mid r \cap q_2^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{s})^{\pm 1} \subseteq r\} & \mathcal{B}_3 = \{r \mid r \cap q_5^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{f}^{-1})^{\pm 1} \subseteq r\} \\
\mathcal{A}_4 = \{r \mid r \cap q_3^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{s})^{\pm 1} \subseteq r\} & \mathcal{B}_4 = \{r \mid r \cap q_6^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{f}^{-1})^{\pm 1} \subseteq r\}
\end{array}$$

$$\mathcal{S}^* = \left\{ r \,\middle|\, \begin{array}{l} 1)\ r \cap q_1^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{f}^{-1})^{\pm 1} \subseteq r, \\ 2)\ r \cap (\mathsf{ss}^{-1}) \neq \emptyset \Rightarrow (\equiv) \subseteq r \end{array} \right\} \quad \mathcal{E}^* = \left\{ r \,\middle|\, \begin{array}{l} 1)\ r \cap q_4^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{s})^{\pm 1} \subseteq r, \\ 2)\ r \cap (\mathsf{ff}^{-1}) \neq \emptyset \Rightarrow (\equiv) \subseteq r \end{array} \right\}$$

$$\mathcal{A}_{\equiv} = \{r \mid r \neq \emptyset \Rightarrow (\equiv) \subseteq r\}$$

$$\mathcal{H} = \left\{ r \,\middle|\, \begin{array}{l} 1)\ r \cap (\mathsf{os})^{\pm 1} \neq \emptyset \ \&\ r \cap (\mathsf{o}^{-1}\mathsf{f})^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{d})^{\pm 1} \subseteq r, \text{ and} \\ 2)\ r \cap (\mathsf{ds})^{\pm 1} \neq \emptyset \ \&\ r \cap (\mathsf{d}^{-1}\mathsf{f}^{-1})^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{o})^{\pm 1} \subseteq r, \text{ and} \\ 3)\ r \cap (\mathsf{pm})^{\pm 1} \neq \emptyset \ \&\ r \not\subseteq (\mathsf{pm})^{\pm 1} \Rightarrow (\mathsf{o})^{\pm 1} \subseteq r \end{array} \right\}$$

$$\mathcal{C}_{\mathsf{o}} = \{r \mid r \neq \emptyset \Rightarrow (\mathsf{oo}^{-1}) \subseteq r\}$$

$$\mathcal{C}_{\mathsf{m}} = \left\{ r \,\middle|\, \begin{array}{l} 1)\ r \neq \emptyset \Rightarrow (\mathsf{mm}^{-1}\mathsf{ss}^{-1}\mathsf{ff}^{-1}) \subseteq r, \text{ and} \\ 2)\ r \cap (\mathsf{pp}^{-1}\mathsf{oo}^{-1}) \neq \emptyset \Rightarrow (\equiv) \subseteq r \end{array} \right\}$$

$$\mathcal{D}_{\mathsf{s}} = \left\{ r \,\middle|\, \begin{array}{l} 1)\ r \cap (\mathsf{dsf})^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{s})^{\pm 1} \subseteq r, \text{ and} \\ 2)\ r \cap (\mathsf{pp}^{-1}\mathsf{mm}^{-1}\mathsf{oo}^{-1}) \neq \emptyset \Rightarrow (\equiv \mathsf{ss}^{-1}) \subseteq r \end{array} \right\}$$

$$\mathcal{D}_{\mathsf{f}} = \left\{ r \,\middle|\, \begin{array}{l} 1)\ r \cap (\mathsf{dsf})^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{f})^{\pm 1} \subseteq r, \text{ and} \\ 2)\ r \cap (\mathsf{pp}^{-1}\mathsf{mm}^{-1}\mathsf{oo}^{-1}) \neq \emptyset \Rightarrow (\equiv \mathsf{ff}^{-1}) \subseteq r \end{array} \right\}$$

$$\mathcal{D}_{\mathsf{d}} = \left\{ r \,\middle|\, \begin{array}{l} 1)\ r \cap (\mathsf{dsf})^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{d})^{\pm 1} \subseteq r, \text{ and} \\ 2)\ r \cap (\mathsf{pp}^{-1}\mathsf{mm}^{-1}\mathsf{oo}^{-1}) \neq \emptyset \Rightarrow (\equiv \mathsf{dd}^{-1}) \subseteq r \end{array} \right\}$$

$$\mathcal{D}'_{\mathsf{d}} = \left\{ r \,\middle|\, \begin{array}{l} 1)\ r \cap (\mathsf{dsf})^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{d})^{\pm 1} \subseteq r, \text{ and} \\ 2)\ r \cap (\mathsf{pmo})^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{odd}^{-1})^{\pm 1} \subseteq r \end{array} \right\}$$

$$\mathcal{D}''_{\mathsf{d}} = \left\{ r \,\middle|\, \begin{array}{l} 1)\ r \cap (\mathsf{dsf})^{\pm 1} \neq \emptyset \Rightarrow (\mathsf{d})^{\pm 1} \subseteq r, \text{ and} \\ 2)\ r \cap (\mathsf{pp}^{-1}\mathsf{oo}^{-1}) \neq \emptyset \Rightarrow (\mathsf{oo}^{-1}\mathsf{dd}^{-1}) \subseteq r, \text{ and} \\ 3)\ r \cap (\mathsf{pp}^{-1}\mathsf{mm}^{-1}) \neq \emptyset \Rightarrow (\equiv \mathsf{dd}^{-1}) \subseteq r \end{array} \right\}$$

Fig. 1. 25 subalgebras of Allen's algebra.

In this paper we present a complete complexity classification for a more general problem, namely, for $\mathcal{A}$-SAT$(\mathcal{F})$ extended with constraints on the lengths of intervals. Now we define the exact form of constraints on lengths we shall allow.

Definition 2. *Let V be a set of real-valued variables, and α, β linear polynomials over V with rational coefficients. A* linear relation *over V is an expression of the form $\alpha R \beta$, where $R \in \{<, \leq, =, \neq, \geq, >\}$.*

A disjunctive linear relation *(DLR) over V is a disjunction of a nonempty finite set of linear relations. A DLR is said to be* Horn *if and only if at most one of its disjuncts is not of the form $\alpha \neq \beta$.*

The problem of satisfiability *for finite sets D of DLRs, denoted* DLRSAT*, is that of checking whether there exists an assignment f of variables in V to real numbers such that all DLRs in D are satisfied. Such an f is said to be a* model *of D. The satisfiability problem for finite sets of Horn DLRs is denoted* HORNDLRSAT.

Theorem 2 ([23, 24]). *The problem* DLRSAT *is NP-complete and* HORNDLRSAT *is solvable in polynomial time.*

We are interested in how the complexity of a problem depends on the value of parameter $\mathcal{F}$ which, in our case, is a set of qualitative relations. Therefore we shall allow only those constraints on lengths which can be expressed by Horn DLRs. This class of constraints subsumes all forms of constraints on lengths which have been considered in [16, 19].

We can now define the general interval satisfiability problem with constraints on lengths.

Definition 3. *An instance of the problem of* interval satisfiability with constraints on lengths *for a set $\mathcal{F} \subseteq \mathcal{A}$, denoted $\mathcal{A}^l$-SAT$(\mathcal{F})$, is a pair $Q = (I, D)$, where*

- *I is an instance of $\mathcal{A}$-SAT$(\mathcal{F})$ over a set V of variables, and*
- *D is an instance of* HORNDLRSAT *over the set of variables $\{l(v) \mid v \in V\}$.*

The question is whether Q is satisfiable*, i.e., whether there exists a model f of I such that the DLRs in D are satisfied with $l(v)$ equal to the length of $f(v)$ for all $v \in V$.*

Proposition 1. *$\mathcal{A}^l$-SAT$(\mathcal{F}) \in$ NP for every $\mathcal{F} \subseteq \mathcal{A}$.*

Proof. Every instance of $\mathcal{A}^l$-SAT$(\mathcal{F})$ over a set V of variables can be translated in a straightforward way into an instance of DLRSAT over the set $\{v^-, v^+ \mid v \in V\}$ of variables. Now the proposition follows from Theorem 2. $\square$

The complexity of $\mathcal{A}^l$-SAT$(\mathcal{S})$ has already been determined for each subalgebra $\mathcal{S}$ listed in Theorem 1.

Proposition 2 ([3]). *The problem $\mathcal{A}^l$-SAT$(\mathcal{S})$ is tractable for $\mathcal{S} \in \{\mathcal{S}_\mathsf{p}, \mathcal{E}_\mathsf{p}, \mathcal{H}\}$ and is NP-complete for the other 15 subalgebras from Theorem 1.*

3 Main Result

Theorem 3. *For any subset $\mathcal{F}$ of $\mathcal{A}$, either $\mathcal{A}^l$-SAT($\mathcal{F}$) is strongly NP-complete or $\mathcal{F}$ is included in $\mathcal{S}$, where $\mathcal{S}$ is one of $\mathcal{S}_\mathsf{p}$, $\mathcal{E}_\mathsf{p}$, $\mathcal{H}$, $\mathcal{C}_\mathsf{o}$, $\mathcal{C}_\mathsf{m}$, $\mathcal{D}_\mathsf{s}$, $\mathcal{D}_\mathsf{f}$, $\mathcal{D}_\mathsf{d}$, $\mathcal{D}'_\mathsf{d}$, $\mathcal{D}''_\mathsf{d}$ (see Fig. 1), for which $\mathcal{A}^l$-SAT($\mathcal{S}$) is tractable.*

In Subsection 3.1, we give polynomial-time algorithms for the 10 subalgebras listed in Theorem 3. Subsection 3.2 contains two NP-completeness results we use to obtain a complete classification of complexity of $\mathcal{A}^l$-SAT($\mathcal{F}$). Finally, in Subsection 3.3, we give a sketch of the classification proof. All correctness, NP-completeness, and classification proofs omitted due to space constraints can be found in the full version of this paper [28].

Strong NP-completeness of the NP-complete cases follows from the fact the the biggest number used in NP-completeness proofs is 5.

3.1 Tractability results

In this subsection we give algorithms for the tractable cases (Fig. 2). Checking that they are polynomial-time is easy and is left to the reader.

We will also assume that algorithms A_i, $1 \le i \le 4$, and the procedure P take instance $Q = (I, D)$ over a set of variables V as an input, and D always contains all constraints of the form $l(v) > 0$, $v \in V$. We will suppose that I does not contain a constraint vrw where $r = \emptyset$. This trivial necessary condition for satisfiability can obviously be checked in polynomial time.

Proposition 3. *The problem $\mathcal{A}^l$-SAT($\mathcal{S}$) is tractable whenever $\mathcal{S}$ is one of $\mathcal{S}_\mathsf{p}$, $\mathcal{E}_\mathsf{p}$, $\mathcal{H}$, $\mathcal{C}_\mathsf{o}$, $\mathcal{C}_\mathsf{m}$, $\mathcal{D}_\mathsf{s}$, $\mathcal{D}_\mathsf{f}$, $\mathcal{D}_\mathsf{d}$, $\mathcal{D}'_\mathsf{d}$, and $\mathcal{D}''_\mathsf{d}$.*

Polynomial-time algorithms solving $\mathcal{A}^l$-SAT($\mathcal{S}$) for $\mathcal{S} \in \{\mathcal{S}_\mathsf{p}, \mathcal{E}_\mathsf{p}, \mathcal{H}\}$ are given in [3]. It is easy to show that an instance $Q = (I, D)$ of $\mathcal{A}^l$-SAT($\mathcal{C}_\mathsf{o}$) is satisfiable if and only if D is satisfiable. Algorithms for the remaining 6 subalgebras are given in Fig. 2.

The following lemma is crucial in the correctness proofs.

Lemma 1 ([21]). *Let D be a satisfiable set of Horn DLRs and let $x_1, \ldots, x_n$ be the variables used in D. If $\tilde{D} = \{x_i \neq x_j \mid D \cup \{x_i \neq x_j\}$ is satisfiable$\}$ then $D \cup \tilde{D}$ is satisfiable.*

Using this lemma we can always divide V into classes such that, in every model of an instance, variables from the same class must be assigned intervals of the same length while any variables from different classes can be assigned intervals of different lengths all at the same time.

Input: instance $Q = (I, D)$ of $\mathcal{A}^l$-SAT$(\mathcal{S})$ with set V of variables

Algorithm A_1 for $\mathcal{S} = \mathcal{C}_{\mathsf{m}}$

1) if D is not satisfiable then reject
2) construct a graph $G = (V, E)$ where $(v, w) \in E$ if and only if
   - $D \cup \{l(v) \neq l(w)\}$ is not satisfiable, and
   - $vrw \in I$ for some r such that $(\equiv) \not\subseteq r$
3) if G is 2-colorable then accept else reject

Procedure P

1) let $D' = D$
2) for each $vrw \in I$ such that $r \subseteq (\mathsf{dsf})$ or $r \subseteq (\mathsf{d}^{-1}\mathsf{s}^{-1}\mathsf{f}^{-1})$, add the constraint $l(v) < l(w)$ or $l(v) > l(w)$, respectively, to D'
3) for each $vrw \in I$ such that $(\equiv) \subseteq r \subseteq (\equiv \mathsf{dsf})$ or $(\equiv) \subseteq r \subseteq (\equiv \mathsf{d}^{-1}\mathsf{s}^{-1}\mathsf{f}^{-1})$, add the constraint $l(v) \leq l(w)$ or $l(v) \geq l(w)$, respectively, to D'
4) if D' is not satisfiable then reject else return (I, D')

Algorithm A_2 for $\mathcal{S} \in \{\mathcal{D}_{\mathsf{s}}, \mathcal{D}_{\mathsf{f}}, \mathcal{D}_{\mathsf{d}}\}$

1) call procedure P
2) accept

Algorithm A_3 for $\mathcal{S} = \mathcal{D}'_{\mathsf{d}}$

1) call procedure P
2) construct a graph $G = (V, E)$ where $(v, w) \in E$ if and only if $D' \cup \{l(v) \neq l(w)\}$ is not satisfiable
3) identify the connected components $S_1, \ldots, S_k$ of G
4) for each S_j, let $I_j = \{vrw \in I \mid v, w \in S_j\}$ and $I'_j = \{v\ r \cap (\equiv \mathsf{oo}^{-1})\ w \mid vrw \in I_j\}$
5) solve I'_j, $1 \leq j \leq k$, as instances of $\mathcal{A}$-SAT$(\mathcal{S}_{\mathsf{o}})$
6) if every I'_j is satisfiable then accept else reject

Algorithm A_4 for $\mathcal{S} = \mathcal{D}''_{\mathsf{d}}$

1) call procedure P
2) construct a graph $G = (V, E)$ where $(v, w) \in E$ if and only if
   - $D' \cup \{l(v) \neq l(w)\}$ is not satisfiable, and
   - $vrw \in I$ for some r such that $(\equiv) \subseteq r \cap (\equiv \mathsf{pp}^{-1}\mathsf{oo}^{-1}\mathsf{mm}^{-1}) \subseteq (\equiv \mathsf{mm}^{-1})$
3) identify the connected components $S_1, \ldots, S_k$ of G
4) for each S_j, let $I_j = \{vrw \in I \mid v, w \in S_j\}$ and $I'_j = \{vrw \mid vrw \in I_j \text{ and } (\equiv) \not\subseteq r\}$
5) if every I'_j is empty then accept else reject

Fig. 2. Polynomial-time algorithms for tractable cases of $\mathcal{A}^l$-SAT.

3.2 NP-completeness results

First let us mention an obvious fact that, for any $\mathcal{F} \subseteq \mathcal{A}$, NP-completeness of $\mathcal{A}$-SAT$(\mathcal{F})$ implies NP-completeness of $\mathcal{A}^l$-SAT$(\mathcal{F})$.

Lemma 2. *Let $r_1, \ldots, r_n \in \mathcal{A}$ be relations such that $\mathcal{A}$-SAT$(\{r_1, \ldots, r_n\})$ is NP-complete.*

1. *If, for every $1 \leq i \leq n$, $r'_i \in \{r_i, r_i \cup (\equiv)\}$ then $\mathcal{A}^l$-SAT$(\{r'_1, \ldots, r'_n\})$ is NP-complete.*
2. *If $\emptyset \neq r_1 \subseteq (\mathsf{pmo})$ and r'_1 is such that $r_1 \subseteq r'_1 \subseteq r_1 \cup (\equiv \mathsf{dsf})$ then $\mathcal{A}^l$-SAT$(\{r'_1, r_2, \ldots, r_n\})$ is NP-complete.*

Lemma 3. *$\mathcal{A}^l$-SAT$(\mathcal{F})$ is NP-complete whenever $\mathcal{F}$ is one of the following sets:*

1. *$\{(\mathsf{oo}^{-1}), (\mathsf{s})\}$, $\{(\mathsf{oss}^{-1}\mathsf{ff}^{-1})\}$, or $\{(\mathsf{sf}), (\mathsf{oo}^{-1}\mathsf{ss}^{-1}\mathsf{ff}^{-1})\}$;*
2. *$\{r\}$ where $(\mathsf{mm}^{-1}) \subseteq r \subseteq (\mathsf{mm}^{-1}\mathsf{dd}^{-1}\mathsf{ss}^{-1}\mathsf{ff}^{-1})$ and $(\mathsf{ss}^{-1}\mathsf{ff}^{-1}) \not\subseteq r$;*
3. *$\{r_1, r_2\}$ where $r_1 \cap (\equiv \mathsf{pp}^{-1}\mathsf{oo}^{-1}\mathsf{mm}^{-1}) = (\mathsf{mm}^{-1}) \subsetneq r_2 \cap (\equiv \mathsf{pp}^{-1}\mathsf{oo}^{-1}\mathsf{mm}^{-1})$ and $(\equiv) \not\subseteq r_2$.*

3.3 Classification of complexity

Now let us briefly discuss the notion of *derivation with lengths* which is the main tool in our classification proofs. This notion is an extension of the notion of derivation in Allen's algebra used in [9].

Suppose $\mathcal{F} \subseteq \mathcal{A}$ and $Q = (I, D)$ is an instance of $\mathcal{A}^l$-SAT$(\mathcal{F})$. Let variables x, y be involved in I. Suppose a relation $r \in \mathcal{A}$ satisfies the following condition: Q is satisfiable if and only if xry. Then we say that r is *derived (with lengths)* from $\mathcal{F}$. It can easily be checked that the problems $\mathcal{A}^l$-SAT$(\mathcal{F})$ and $\mathcal{A}^l$-SAT$(\mathcal{F} \cup \{r\})$ are polynomially equivalent because, in any instance of the second problem, any constraint involving r can be replaced by the set of constraints in Q (introducing fresh variables when needed), and this can be done in polynomial time. It follows that it is sufficient to classify the complexity of problems $\mathcal{A}^l$-SAT$(\mathcal{S})$ where $\mathcal{S}$ is a subalgebra of $\mathcal{A}$ closed under derivations with lengths.

The classification proof has 8 steps. In each step, it is proved that if a subalgebra $\mathcal{S}$ which is closed under derivations with lengths satisfies a certain condition then either $\mathcal{S}$ is contained in one of the 10 tractable subalgebras, or Lemma 2 or 3 can be applied to some $\mathcal{F} \subseteq \mathcal{S}$, or $\mathcal{S}$ satisfies the conditions of one of the previous steps.

We shall say that a relation is *non-trivial* if it is not equal to the empty relation or the relation $(\equiv)$.

Step 1. Suppose $\mathcal{S}$ contains a non-trivial relation r with $r \subseteq (\equiv \mathsf{pp}^{-1}\mathsf{mm}^{-1}\mathsf{oo}^{-1})$. Then either $\mathcal{S}$ is contained in one of $\mathcal{C}_{\mathsf{o}}$, $\mathcal{S}_{\mathsf{p}}$, $\mathcal{E}_{\mathsf{p}}$, and $\mathcal{H}$, or else $\mathcal{A}^l$-SAT$(\mathcal{S})$ is NP-complete.
Step 2. Suppose $\mathcal{S}$ contains a non-trivial relation r such that $r \cap r^{-1} \subseteq (\equiv)$ and neither r nor r^{-1} is contained in $(\equiv \mathsf{dsf})$. Then either $\mathcal{S}$ is contained in one of $\mathcal{C}_{\mathsf{o}}$, $\mathcal{S}_{\mathsf{p}}$, $\mathcal{E}_{\mathsf{p}}$, and $\mathcal{H}$ or else $\mathcal{A}^l$-SAT$(\mathcal{S})$ is NP-complete.

Step 3. If $\mathcal{S}$ contains two non-trivial relations r_1 and r_2 such that $r_1 \cap r_2 \subseteq (\equiv)$ and $r_1, r_2 \subseteq (\equiv \mathsf{dsf})$ then either $\mathcal{S} \subseteq \mathcal{H}$ or else $\mathcal{A}^l$-SAT$(\mathcal{S})$ is NP-complete.
Step 4. If $\mathcal{S}$ contains two non-trivial symmetric relations r_1 and r_2 such that $r_1 \cap r_2 \subseteq (\equiv)$ then either $\mathcal{S}$ is contained in one of $\mathcal{S}_{\mathsf{p}}$, $\mathcal{E}_{\mathsf{p}}$, $\mathcal{H}$, or else $\mathcal{A}^l$-SAT$(\mathcal{S})$ is NP-complete.
Step 5. If $(\mathsf{s}) \in \mathcal{S}$ or $(\mathsf{f}) \in \mathcal{S}$ then either $\mathcal{S}$ is contained in one of the 10 subalgebras listed in Theorem 3 or $\mathcal{A}^l$-SAT$(\mathcal{S})$ is NP-complete.
Step 6. If $(\mathsf{sf}) \in \mathcal{S}$ then either $\mathcal{S} \subseteq \mathcal{D}_{\mathsf{s}}$ or $\mathcal{S} \subseteq \mathcal{D}_{\mathsf{f}}$ or else $\mathcal{A}^l$-SAT$(\mathcal{S})$ is NP-complete.
Step 7. If there is $r \in \mathcal{S}$ such that $(\mathsf{d}) \subseteq r \subseteq (\mathsf{dsf})$ then either $\mathcal{S}$ is contained in one of the 10 subalgebras listed in Theorem 3 or else $\mathcal{A}^l$-SAT$(\mathcal{S})$ is NP-complete.
Step 8. If there is a symmetric non-trivial relation $r' \in \mathcal{S}$ such that every non-trivial $r \in \mathcal{S}$ satisfies $r' \subseteq r$ then either $\mathcal{S}$ is contained in one of the 10 tractable subalgebras or $\mathcal{A}^l$-SAT$(\mathcal{S})$ is NP-complete.

Note that the assumptions of these 8 steps are easily shown to be exhaustive. Indeed, if there is $r \in \mathcal{S}$ such that $r \cap r^{-1} \subseteq (\equiv)$ then either Step 2 or one of Steps 5-7 applies because, due to closedness under derivations with lengths, a subalgebra containing $r \cup (\equiv)$ where $r \subseteq (\mathsf{dsf})$ also contains r itself. Otherwise Step 4 or Step 8 applies, since $r \cap r^{-1} \in \mathcal{S}$ for every $r \in \mathcal{S}$.

Classification is complete. Theorem 3 is proved. □

References

[1] Montanari, U.: Networks of constraints: Fundamental properties and applications to picture processing. Information Sciences **7** (1974) 95–132
[2] Garey, M., Johnson, D.: Computers and Intractability: A Guide to the Theory of NP-Completeness. Freeman, New York (1979)
[3] Angelsmark, O., Jonsson, P.: Some observations on durations, scheduling and Allen's algebra. In: Proceedings of the 6th Conference on Constraint Programming (CP'00). Volume 1894 of Lecture Notes in Computer Science., Springer-Verlag (2000) 484–488
[4] Jeavons, P.: On the algebraic structure of combinatorial problems. Theoretical Computer Science **200** (1998) 185–204
[5] Schaefer, T.: The complexity of satisfiability problems. In: Proceedings of the 10th Symposium on the Theory of Computing (STOC'78), New-York, ACM Press (1978) 216–226
[6] Hell, P., Nešetřil, J.: On the complexity of H-coloring. Journal of Combinatorial Theory, Ser. B **48** (1990) 92–110
[7] Allen, J.: Maintaining knowledge about temporal intervals. Communications of the ACM **26** (1983) 832–843
[8] Golumbic, M., Shamir, R.: Complexity and algorithms for reasoning about time: A graph-theoretic approach. Journal of the ACM **40** (1993) 1108–1133
[9] Krokhin, A., Jeavons, P., Jonsson, P.: Reasoning about temporal relations: Tractable subalgebras of Allen's interval algebra. Technical Report RR-01-12, Oxford University Computing Laboratory (2001) Submitted for publication.

[10] Nebel, B., Bürckert, H.J.: Reasoning about temporal relations: A maximal tractable subclass of Allen's interval algebra. Journal of the ACM **42** (1995) 43–66

[11] Goldberg, P., Golumbic, M., Kaplan, H., Shamir, R.: Four strikes against physical mapping of DNA. Journal of Computational Biology **2** (1995) 139–152

[12] Golumbic, M., Kaplan, H., Shamir, R.: On the complexity of DNA physical mapping. Advances in Applied Mathematics **15** (1994) 251–261

[13] Karp, R.: Mapping the genome: some combinatorial problems arising in molecular biology. In: Proceedings of the 25th Symposium on the Theory of Computing (STOC'93), New-York, ACM Press (1993) 278–285

[14] Hsu, W.L.: A simple test for interval graphs. In Hsu, W.L., Lee, R., eds.: Proceedings of the 18th International Workshop on Graph-Theoretic Concepts in Computer Science. Volume 657 of Lecture Notes in Computer Science., Springer (1992) 11–16

[15] Corneil, D., Kim, H., Natarajan, S., Olariu, S., Sprague, A.: Simple linear time recognition of unit interval graphs. Information Processing Letters **55** (1995) 99–104

[16] Pe'er, I., Shamir, R.: Realizing interval graphs with size and distance constraints. SIAM Journal on Discrete Mathematics **10** (1997) 662–687

[17] Gabor, C., Supowit, K., Hsu, W.: Recognizing circle graphs in polynomial time. Journal of the ACM **36** (1989) 435–473

[18] Golumbic, M., Kaplan, H., Shamir, R.: Graph sandwich problems. Journal of Algorithms **19** (1995) 449–473

[19] Pe'er, I., Shamir, R.: Satisfiability problems on intervals and unit intervals. Theoretical Computer Science **175** (1997) 349–372

[20] Webber, A.: Proof of the interval satisfiability conjecture. Annals of Mathematics and Artificial Intelligence **15** (1995) 231–238

[21] Drakengren, T., Jonsson, P.: Eight maximal tractable subclasses of Allen's algebra with metric time. Journal of Artificial Intelligence Research **7** (1997) 25–45

[22] Drakengren, T., Jonsson, P.: A complete classification of tractability in Allen's algebra relative to subsets of basic relations. Artificial Intelligence **106** (1998) 205–219

[23] Jonsson, P., Bäckström, C.: A unifying approach to temporal constraint reasoning. Artificial Intelligence **102** (1998) 143–155

[24] Koubarakis, M.: Tractable disjunctions of linear constraints: basic results and applications to temporal reasoning. Theoretical Computer Science **266** (2001) 311–339

[25] Fortune, S., Hopcroft, J., Wyllie, J.: The directed subgraph homeomorphism problem. Theoretical Computer Science **10** (1980) 111–121

[26] Creignou, N., Khanna, S., Sudan, M.: Complexity Classification of Boolean Constraint Satisfaction Problems. Volume 7 of SIAM Monographs on Discrete Mathematics and Applications. (2001)

[27] Kirousis, L., Kolaitis, P.: The complexity of minimal satisfiability problems. In: Proceedings of the 18th Symposium on Theoretical Aspects of Computer Science (STACS 2001). Volume 2010 of Lecture Notes in Computer Science., Springer (2001) 407–418

[28] Krokhin, A., Jeavons, P., Jonsson, P.: The complexity of constraints on intervals and lengths. Technical Report TR01-077, Electronic Colloquim on Computational Complexity (2001) (`http://www.eccc.uni-trier.de/eccc`).