Journal of Computer and System Sciences 82 (2016) 912–928

Contents lists available at ScienceDirect

Journal of Computer and System Sciences

www.elsevier.com/locate/jcss

Constraint satisfaction and semilinear expansions of addition over the rationals and the reals

Peter Jonsson [a,1], Johan Thapper [b,*]

[a] Department of Computer and Information Science, Linköpings universitet, SE-58183 Linköping, Sweden
[b] Laboratoire d'Informatique Gaspard-Monge, Université Paris-Est Marne-la-Vallée, 5 boulevard Descartes, F-77420 Champs-sur-Marne, France

ARTICLE INFO

Article history:
Received 18 May 2015
Received in revised form 22 February 2016
Accepted 7 March 2016
Available online 17 March 2016

Keywords:
Constraint satisfaction problems
Semilinear sets
Algorithms
Computational complexity

ABSTRACT

A semilinear relation is a finite union of finite intersections of open and closed half-spaces over, for instance, the reals, the rationals, or the integers. Semilinear relations have been studied in connection with algebraic geometry, automata theory, and spatiotemporal reasoning. We consider semilinear relations over the rationals and the reals. Under this assumption, the computational complexity of the constraint satisfaction problem (CSP) is known for all finite sets containing $R_+ = \{(x, y, z) \mid x + y = z\}$, $\leq$, and $\{1\}$. These problems correspond to expansions of the linear programming feasibility problem. We generalise this result and fully determine the complexity for all finite sets of semilinear relations containing R_+. This is accomplished in part by introducing an algorithm, based on computing affine hulls, which solves a new class of semilinear CSPs in polynomial time. We further analyse the complexity of linear optimisation over the solution set and the existence of integer solutions.

1. Introduction

We work over a ground set (or domain) X, which will either be the rationals, $\mathbb{Q}$, or the reals, $\mathbb{R}$. We say that a relation $R \subseteq X^k$ is *semilinear* if it can be represented as a finite union of finite intersections of open and closed half-spaces in X^k. Alternatively, R is semilinear if it is first-order definable in $\{R_+, \leq, \{1\}\}$ where $R_+ = \{(x, y, z) \in X^3 \mid x + y = z\}$ [1]. Semilinear relations appear in many different contexts within mathematics and computer science: they are, for instance, frequently encountered in algebraic geometry, automata theory, spatiotemporal reasoning, and computer algebra. Semilinear relations have also attained a fair amount of attention in connection with *constraint satisfaction problems* (CSPs). In a CSP, we are given a set of variables, a (finite or infinite) domain of values, and a finite set of constraints. The question is whether or not we can assign values to the variables so that all constraints are satisfied. From a complexity theoretical viewpoint, solving general constraint satisfaction problems is obviously a hard problem. Various ways of refining the problem can be adopted to allow a more meaningful analysis. A common refinement is that of introducing a *constraint language*; a finite set Γ of allowed relations. One then considers the problem CSP(Γ) in which all constraint in the input must be members of Γ. This parameterisation of constraint satisfaction problems has proved to be very fruitful for CSPs over both finite and infinite

* Corresponding author.
E-mail addresses: peter.jonsson@liu.se (P. Jonsson), thapper@u-pem.fr (J. Thapper).
[1] The first author was partially supported by the Swedish Research Council (VR) under grant 621-2012-3239.

http://dx.doi.org/10.1016/j.jcss.2016.03.002

domains. Since Γ is finite, the computational complexity of such a problem does not depend on the actual representation of the constraints.

The complexity of finite-domain CSPs has been studied for a long time and a powerful algebraic toolkit has gradually formed [2]. Much of this work has been devoted to the *Feder–Vardi conjecture* [3,4] which posits that every finite-domain CSP is either polynomial-time solvable or NP-complete. Infinite-domain CSPs, on the other hand, constitute a much more diverse set of problems. In fact, *every* computational problem is polynomial-time equivalent to an infinite-domain CSP [5]. Obtaining a full understanding of their computational complexity is thus out of the question, and some further restriction is necessary. In this article, this restriction will be to study semilinear relations and constraint languages.

A relation $R \subseteq X^k$ is said to be *essentially convex* if for all $p, q \in R$ there are only finitely many points on the line segment between p and q that are not in R. A constraint language Γ is said to be essentially convex if every relation in Γ is essentially convex. The main motivation for this study is the following result:

Theorem 1. *(See Bodirsky et al. [6].) Let Γ be a finite set of semilinear relations over $\mathbb{Q}$ or $\mathbb{R}$ such that $\{R_+, \leq, \{1\}\} \subseteq \Gamma$. Then,*

1. *CSP(Γ) is polynomial-time solvable if Γ is essentially convex, and NP-complete otherwise; and*
2. *the problem of optimising a linear polynomial over the solution set of CSP(Γ) is polynomial-time solvable if and only if CSP(Γ) is polynomial-time solvable (and NP-hard otherwise).*

One may suspect that there are semilinear constraint languages Γ such that CSP$(\Gamma) \in P$ but Γ is not essentially convex. This is indeed true and we identify two such cases. In the first case, we consider relations with large "cavities". It is not surprising that the algorithm for essentially convex relations (and the ideas behind it) cannot be applied in the presence of such highly non-convex relations. Thus, we introduce a new algorithm which solves CSPs of this type in polynomial time. It is based on computing affine hulls and the idea of improving an easily representable upper bound on the solution space by looking at one constraint at a time; a form of "local consistency" method. In the second case, we consider relations R that are not necessarily essentially convex, but look essentially convex when viewed form the origin. That is, any points p and q that witness a not essentially convex relation lie on a line that passes outside of the origin. We show that we can remove all such holes from R to find an equivalent constraint language that is essentially convex, and thereby solve the problem in polynomial time.

Combining these algorithmic results with matching NP-hardness results and the fact that CSP(Γ) is always in NP for a semilinear constraint language Γ (cf. Theorem 5.2 in Bodirsky et al. [6]) yields a dichotomy:

Theorem 2. *Let Γ be a finite set of semilinear constraints that contains R_+. Then, CSP(Γ) is either in P or NP-complete.*

Our result immediately generalises the first part of Theorem 1. It also generalises another result by Bodirsky et al. [7] concerning expansions of $\{R_+\}$ with relations that are first-order definable in $\{R_+\}$. One may note that this class of relations is a severely restricted subset of the semilinear relations since it admits quantifier elimination over the structure $\{+, \{0\}\}$, where $+$ denotes the binary addition function. This follows from the more general fact that the first-order theory of torsion-free divisible abelian groups admits quantifier elimination (see e.g. Theorem 3.1.9 in [8]). One may thus alternatively view relations that are first-order definable in $\{R_+\}$ as finite unions of sets defined by homogeneous linear systems of equations.

We continue by generalising the second part of Theorem 1, too: if Γ is semilinear and contains $\{R_+, \{1\}\}$, then the problem of optimising a linear polynomial over the solution set of CSP(Γ) is polynomial-time solvable if and only if CSP(Γ) is polynomial-time solvable (and NP-hard otherwise). We also study the problem of finding integer solutions to CSP(Γ) for certain semilinear constraint languages Γ. Here, we obtain some partial results but a full classification remains elusive. Our results shed some light on the *scalability property* introduced by Jonsson and Lööw [9].

This article has the following structure. We begin by formally defining constraint satisfaction problems and semilinear relations together with some terminology and minor results in Section 2. The algorithms and tractability results are presented in Section 3 while the hardness results can be found in Section 4. By combining the results from Sections 3 and 4, we prove Theorem 2 in Section 5. We partially generalise Theorem 2 to optimisation problems in Section 6, and we study the problem of finding integer solutions in Section 7. Finally, we discuss some obstacles to further generalisations in Section 8. This article is a revised and extended version of a conference paper [10].

2. Preliminaries

2.1. Constraint satisfaction problems

Let $\Gamma = \{R_1, \ldots, R_n\}$ be a finite set of finitary relations over some domain X (which will usually be infinite). We refer to Γ as a *constraint language*. In order to avoid some uninteresting trivial cases, we will assume that all constraint languages are non-empty and contain non-empty relations only.

A first-order formula is called *primitive positive* if it is of the form $\exists x_1, \ldots, x_n \,.\, \psi_1 \wedge \cdots \wedge \psi_m$, where each ψ_i is an atomic formula, i.e., either $x = y$ or $R(x_{i_1}, \ldots, x_{i_k})$ with R the relation symbol for a k-ary relation from Γ. We call such a formula