Available online at www.sciencedirect.com
ScienceDirect
Information and Computation 205 (2007) 651–678
Information and Computation
www.elsevier.com/locate/ic

Towards a dichotomy theorem for the counting constraint satisfaction problem

Andrei A. Bulatov [a,*], Víctor Dalmau [b]

[a] *School of Computing Science, Simon Fraser University, 8888 University Drive, Burnaby, Canada V5A 1S6*
[b] *Department de Tecnologia, Universitat Pompeu Fabra, Barcelona, Spain*

Received 27 July 2005; revised 27 September 2006
Available online 15 November 2006

Abstract

The counting constraint satisfaction problem (#CSP) can be expressed as follows: given a set of variables, a set of values that can be taken by the variables, and a set of constraints specifying some restrictions on the values that can be taken simultaneously by some variables, determine the number of assignments of values to variables that satisfy all the constraints. The #CSP provides a general framework for numerous counting combinatorial problems including counting satisfying assignments to a propositional formula, counting graph homomorphisms, graph reliability and many others. This problem can be parametrized by the set of relations that may appear in a constraint. In this paper we start a systematic study of subclasses of the #CSP restricted in this way. The ultimate goal of this investigation is to distinguish those restricted subclasses of the #CSP which are solvable in polynomial time from those which are not. We show that the complexity of any restricted #CSP class on a finite domain can be deduced from the properties of polymorphisms of the allowed constraints, similar to that for the decision constraint satisfaction problem. Then we prove that if a subclass of the #CSP is solvable in polynomial time, then constraints allowed by the class satisfy some very restrictive condition: they need to have a Mal'tsev polymorphism, that is a ternary operation $m(x, y, z)$ such that $m(x, y, y) = m(y, y, x) = x$. This condition uniformly explains many existing complexity results for particular cases of the #CSP, including the dichotomy results for the problem of counting graph homomorphisms, and it allows us to obtain new results.

* Corresponding author. Fax: +1 604 291 3045.
E-mail addresses: abulatov@cs.sfu.ca (A.A. Bulatov), victor.dalmau@upf.edu (V. Dalmau).

0890-5401/$ - see front matter
doi:10.1016/j.ic.2006.09.005

Keywords: Constraint satisfaction problem; Counting problems; Complexity

1. Introduction

In a counting combinatorial problem the objective is to find the number of feasible solutions to a certain search problem. Similar to its decision counterpart, the Counting Constraint Satisfaction Problem (#CSP) can be used to provide a generic framework for numerous counting combinatorial problems that arise frequently in a wide range of areas from logic, graph theory, and artificial intelligence [4,13,21,26,33,41,45,51,52,55,56], to statistical physics [3,11,39].

The prototypical counting problem, #SAT, i.e., the problem of counting the number of assignments that satisfy a CNF formula, constitutes an important particular case of the #CSP. Since the pionnering papers of Valiant [55,56] the computational complexity of counting satisfying assignments to propositional formulas of various types [13,41,51,52,55,56] has been intensively investigated. In particular, it has been found that #SAT is much more computationally demanding than its decision counterpart SAT, and is #P-complete even for Horn or monotone formulas, and even when the size of clauses and the number of occurrences of a variable in the formula are extremely limited. In [13], Creignou and Hermann obtained a dichotomy theorem for #SAT, similar to that of Schaefer [53] for (decision) SAT.

The formalism of constraint networks introduced by Montanari [43] provides a natural generalization of propositional formulas to domains with more than 2 elements. A constraint network is given by a collection of variables, a domain, and a family of constraints where a constraint is a pair given by a list of variables, called the *scope*, and a relation indicating the valid combinations of values for the variables in the scope. The problem of deciding whether there exists a solution to a constraint network, i.e., an assignment of values to variables satisfying all the constraints, is known as the constraint satisfaction problem (CSP). This problem received considerable attention in theoretical computer science and it also constitutes one of the major lines of research in artificial intelligence.

The class of counting constraint satisfaction problems is defined as the counting version of the CSP, i.e., the problem of finding the number of solutions to a constraint network. This problem can also be reformulated as (1) the problem of finding the number of models of a conjunctive formula, as (2) the problem of counting the number of homomorphisms between two finite relational structures **A** and **B**, and also as (3) the problem of computing the size (number of tuples) of the evaluation $Q(D)$ of a conjuntive query (without projection) Q on a database D; see [24,37].

Further examples of combinatorial problems expressible in a natural way in #CSP terms include problems from propositional logic [13,52], classical combinatorial problems such as #Clique, Graph Reliability, Antichain, Permanent [41,51,55,56], counting graph homomorphisms, and many others [4,21,26,33].

A particular case of the counting graph homomorphisms problem, the class of #H-Coloring problems, attracts a special attention. In a #H-Coloring problem the goal is to count the number of homomorphisms from a graph G (the input) to a fixed graph H. Recently, Dyer and Greenhill [21] proved that, for every undirected graph H, its associated #H-Coloring problem is either in FP or #P-complete (even when restricted to graphs of bounded degree) and they have also provided a complete characterization of the tractable problems. This result has been extended to the counting