Complexity and Approximability of Quantified and Stochastic Constraint Satisfaction Problems

Harry B. Hunt III [1]

Department of Computer Science, University at Albany – SUNY, Albany, NY 12222, USA. Email: `hunt@cs.albany.edu`

Madhav V. Marathe [2]

Basic and Applied Simulation Sciences, Los Alamos National Laboratory, P.O. Box 1663, MS M997, Los Alamos, NM 87545, USA. Email: `marathe@lanl.gov`

Richard E. Stearns [3]

Department of Computer Science, University at Albany – SUNY, Albany, NY 12222, USA. Email: `res@cs.albany.edu`

Abstract

Let D be an arbitrary (*not* necessarily finite) nonempty set, let C be a finite set of constant symbols denoting arbitrary elements of D, and let $\mathbf{S}$ and $\mathbf{T}$ be an arbitrary finite set of finite-arity relations on D. We denote the problem of determining the satisfiability of finite conjunctions of relations in $\mathbf{S}$ applied to variables (to variables and symbols in C) by SAT($\mathbf{S}$) (by $\mathrm{SAT}_C(\mathbf{S})$.) Here, we study *simultaneously* the complexity of decision, counting, maximization and approximate maximization problems, for unquantified, quantified and stochastically quantified formulas.

We present simple yet general techniques to characterize *simultaneously*, the complexity or efficient approximability of a number of versions/variants of the problems SAT($\mathbf{S}$), Q-SAT($\mathbf{S}$), S-SAT($\mathbf{S}$),MAX-Q-SAT($\mathbf{S}$) etc., for many different such D, C, S, T. These versions/variants include decision, counting, maximization and approximate maximization problems, for unquantified, quantified and stochastically quantified formulas. Our unified approach is based on the following two basic concepts: (i) strongly-local replacements/reductions and (ii) relational/algebraic representability.

Some of the results extend the earlier results in [Pa85,LMP99,CF+93,CF+94] Our techniques and results reported here also provide *significant steps towards obtaining dichotomy theorems*, for a number of the problems above, including the problems MAX-Q-SAT($\mathbf{S}$), and MAX-S-SAT($\mathbf{S}$). The discovery of such dichotomy theorems, for unquantified formulas, has received significant recent attention in the literature [CF+93,CF+94,Cr95,KSW97].

Preprint submitted to Elsevier Science *29 May 2001*

Key words: **NP**-hardness, Approximation Algorithms, **PSPACE**-hardness, Quantified and Stochastic Constraint Satisfaction Problems.

1 Introduction and motivation

Over the past thirty years, researchers in theoretical computer science, AI, operations research, and computational algebra have studied versions/variants of CNF satisfiability, constraint satisfaction problems, and algebraic satisfiability problems (i.e. the problems $\mathrm{SAT}(F)$ where F is an algebraic structure). Two important reasons for the extensive research are the following:
A. Versions/variants of these problems are widely applicable in modeling both real-life and abstract mathematical problems.
B. These problems, especially versions/variants of the problem 3SAT and Q-3SAT, have played fundamental roles in the development of discrete complexity theory, providing prototypical complete problems for various complexity classes.
The results, concepts, and techniques reported here are relevant to the following topical areas of this ongoing research:
1. the complexity of 3SAT and more generally Boolean constraint satisfaction problems, e.g. [Sc78,GJ79,FV93,JCG97,CJ+00]
2. the development of dichotomy type results for decision and optimization versions of Boolean constraint satisfaction problems, e.g. [HSM94,HSM01a,Cr95,KSW97,MH+94,LMP99],
3. the complexity and (non)-approximability of **PSPACE**-hard quantified and stochastic Boolean satisfiability problems, e.g., [FY79,Pa94,Pa85,CF+93,CF+94,MH+94,HSM94,LMS96,LMP99],
4. the complexity of decision and optimization problems, when instances are succinctly specified, e.g. [Le83,LW92,Or82,MH+94], and
5. the complexity of solving systems of equations on various algebraic structures, e.g. [AC+98,Le83,GJ79,IM83,LW87,AB88,HSM01b].
Here combining these lines of research we study the complexity and efficient approximability of quantified and stochastic constraint satisfaction problems. The research program also spans complexity and approximability of solving quantified and stochastically quantified systems of equations on various algebraic structures – due to space limitations we will not discuss these topics in detail. We recall preliminary definitions that will be helpful in understand-

[1] Some of the work done while the author was visiting the Basic and Applied Simulation Sciences Group at the Los Alamos National Laboratory. supported by the Department of Energy under Contract W-7405-ENG-36. Also partially supported by NSF Grant CCR94-06611.
[2] Supported by Department of Energy under Contract W-7405-ENG-36.
[3] Research supported in part by NSF Grant CCR94-06611.

ing our results. let D be an arbitrary (*not* necessarily finite) set of cardinality ≥ 2; C is a finite set of constant symbols denoting distinct elements of D; and $\mathbf{S}$,$\mathbf{T}$ are finite sets of finite-arity relations on D. An $\mathbf{S}$-*clause* (a *constant-free* $\mathbf{S}$-*clause*) is a relation in $\mathbf{S}$ applied to variables and constants in C (to variables, *only*.) An $\mathbf{S}$-*formula* (a *constant-free* $\mathbf{S}$-*formula*) is a finite nonempty conjunction of $\mathbf{S}$-clauses (of constant-free $\mathbf{S}$-clauses.) By $Rep(\mathbf{S})$ (by $Rep_C(\mathbf{S})$), we mean the set of all finite-arity relations on D denoted by existentially-quantified conjunctions of constant-free (of arbitrary) $\mathbf{S}$-clauses. Let D_k-Relations denote the set of all finite arity relations over a domain D of cardinality k. We use $Rep(\mathbf{S}) = D_k$-**RELATIONS** to mean that all finite arity relations on D_k can be equivalently represented as finite existentially-quantified conjunctions of relations in $\mathbf{S}$ applied to variables (to variables and constant symbols in C). An $\mathbf{S}$-formula is *satisfiable* if all of its clauses are *simultaneously* satisfiable. *Quantified* and (when D is *finite*) *stochastically-quantified* $\mathbf{S}$-formulas and their *satisfiability* problems are defined as generally assumed in the literature. Consequently, a quantified $\mathbf{S}$-formula is *satisfiable* **iff** it has at least one *proof-tree*. We follow [Pa85,LMP99], for the definition of the *random* quantifier. The difference being that [Pa85] only allows *existential* and *random* quantifiers in stochastic formulas, while [LMP99] *also* allows *universal* quantifiers. (Both [Pa85,LMP99] only consider stochastically-quantified Boolean formulas. Thus Papadimitriou [Pa85] introduced the *stochastic satisfiability* problem defined as follows: A problem instance is assumed to be of the form–

$$F = \mathbf{R}x_1 \exists x_2 \cdots \mathbf{R}x_{n-1} \exists x_n (\mathbf{E}(F(x_1, \ldots, x_n)) \geq 1/2)),$$

where $\mathbf{R}$ is the *random quantifier–there is a random assignment of truth-values to the variable with equiprobable values of* 1 *and* 0, $\mathbf{E}$ denotes *expectation*, and f is a 3-CNF formula. Under the *equiprobable* assumption, such a formula F is *satisfiable* **iff** there is a "proof-tree" for F such that at least 1/2 of its leaves evaluate to *True*. We leave the discussion of the extensions needed to Papadimitriou's definition to define the STOCHASTIC SATISFIABILITY problems of [LMP99] to the full paper (or see [LMP99].) Finally as observed above by varying D, C, $\mathbf{S}$, the allowed quantifiers, etc., a large number of problems in computer science, operations research, AI, etc., can be expressed *both* naturally and directly in terms of $\mathbf{S}$-, quantified $\mathbf{S}$-,or stochastically-quantified $\mathbf{S}$-formulas. We defined the problems SAT($\mathbf{S}$), SAT$_C$($\mathbf{S}$), ..., S-SAT$_C$($\mathbf{S}$) above. We also consider the following optimization versions of these decision problems:

1. the problems MAX-SAT($\mathbf{S}$), MAX-SAT$_C$($\mathbf{S}$), MAX-Q-SAT($\mathbf{S}$), MAX-Q-SAT$_C$($\mathbf{S}$), MAX-S-SAT($\mathbf{S}$), MAX-S-SAT$_C$($\mathbf{S}$), as defined in [PY91,CF+93,CF+94];

2. the problems MAX-NSF-SAT($\mathbf{S}$), MAX-NSF-Q-SAT($\mathbf{S}$), and MAX-NSF- S-SAT($\mathbf{S}$), that entail finding the maximum number of *simultaneously* satisfiable formulas, quantified formulas, and stochastically-quantified formulas (this last so that the probability of each simultaneously-satisfied formula is greater than some given *fixed* rational threshold Θ $(0 < \Theta \leq 1)$) in a given finite sequence of such formulas;

3. the problems MAX-Q-FORMULA-SAT($\mathbf{S}$) and MAX-S-FORMULA-SAT($\mathbf{S}$) defined as in [CF+93,HSM94]; and

4. (for finite domains D) the problems MAX-SATISFYING-ASSIGN-SAT($\mathbf{S}$) and MAX-DONES-SAT($\mathbf{S}$) as defined for 3CNF formulas in [Kr88,PR93] [4] and their extensions to quantified and stochastically-quantified formulas with *free* variables.

Our main result can be viewed as *relative complexity* result and can be stated roughly as follows: Let $\mathbf{S}$ and $\mathbf{T}$ be defined over **finite domains** (not necessarily the same). Let $\mathbf{S} \subset Rep(\mathbf{T})$. Then, the problem SAT($\mathbf{S}$) is reducible to the problem SAT($\mathbf{T}$) by a reduction $R(\mathbf{S},\mathbf{T})$ such that:

(1) The reduction $R(\mathbf{S},\mathbf{T})$ is a 1-*strongly-local-replacement-reduction*. Consequently, it can be translated into $O(n \cdot logn)$ time- and linear size-bounded reductions $R^Q(\mathbf{S},\mathbf{T})$ and $R^S(\mathbf{S},\mathbf{T})$ of the problems Q-SAT($\mathbf{S}$) and S-SAT($\mathbf{S}$) to the problems Q-SAT($\mathbf{T}$) and S-SAT($\mathbf{T}$), respectively. These last two reductions preserve a number of structural properties of instances. In addition such a reduction $R(\mathbf{S},\mathbf{T})$ can be constructed efficiently from any *representation* of $S \in \mathbf{S}$ by relations in $\mathbf{T}$.

(2) The reduction $R^Q(\mathbf{S},\mathbf{T})$ and $R^S(\mathbf{S},\mathbf{T})$ preserves a number of global properties of *proof* and *partial-proof-trees* including the ratio of true leaves to leaves. It can be modified efficiently into an $O(n \cdot logn)$ time- and linear size-bounded A-reduction of the problems MAX-NSF-Q-SAT($\mathbf{S}$), MAX-Q-FORMULA-SAT($\mathbf{S}$) MAX-NSF-S-SAT($\mathbf{S}$) and MAX-S-FORMULA-SAT($\mathbf{S}$) into the problems MAX-NSF-Q-SAT($\mathbf{T}$), MAX-Q-FORMULA-SAT($\mathbf{T}$), MAX-NSF-S-SAT($\mathbf{T}$) and MAX-S-FORMULA-SAT($\mathbf{T}$), respectively. In addition, we have identified several natural sufficient conditions on a representation of $\mathbf{S}$ by $\mathbf{T}$, for the corresponding reductions $R(\mathbf{S},\mathbf{T})$ and $R^Q(\mathbf{S},\mathbf{T})$ to be *parsimonious*, L, preserve planarity of instances, etc.

Important Note: For clarity, we have stated the above results for **finite domains**. In case of quantified formulas, the above results can be easily extended when the underlying domains are infinite. Similar extension requires some care for stochastic formulas due to the technical issues of defining the right probability measure.

Our potentially most important contributions are the identification and subsequent formalization of the few basic concepts and the derivation of

[4] A MAX-Satisfying-Assignment of a formula is a satisfying assignment to the variables of the formula, that is lexicographically maximum: hence the restriction to finite domains D. A MAX-DONES-satisfying assignment is a satisfying assignment with the maximum number of designated variables set equal to 1.

their complexity-theoretic properties, that suffice to prove these results naturally, fairly directly, and *simultaneously.* In the companion papers [HSM01a,HSM01b], we demonstrate how these concepts can be used to characterize the complexity and efficient approximability of versions/variants of determining the satisfiability of *unquantified* systems of equations and/or constraints on many algebraic structures. Below, we briefly discuss how our results/techniques/constructions, for versions/variants of the algebraic satisfaction problems SAT(F), can be extended to apply to *quantified* and *stochastically-quantified* algebraic satisfaction problems.

Example 1: The generalized CNF satisfiability problems SAT(S) and SAT_c(S) generalize the problems 3SAT, 1-3SAT , NAE-3SAT, etc. in [GJ79]. For example, let $EO(x, y, z)$ be the ternary logical relation given by $\{(1,0,0),(0,1,0),(0,0,1)\}$. Then, the problem EXACTLY-1-IN-EX3-MONO-SAT is the same as the problem SAT({EO}). An instance of the above problem might consist of the set of variables x, y, z, w and the formula $F = EO(x,y,z) \wedge EO(x,y,w) \wedge EO(x,w,z)$. It is easy to see that F is satisfiable by setting $x = 1$ and setting all other variables to 0. Let f be an S-formula with m clauses and n_i literals in clause i, $1 \leq i \leq m$. The size of f denoted by $size(f)$ is given by $O(\sum_{i=1}^{i=m} n_i)$. Let C be a set of clauses defined over a set of variables V. We will use $F(C, V)$ to denote the formula obtained by the conjunction of clauses in C. By appropriately defining unary, binary and ternary versions of the relation EO, it is possible to define 1-3-SAT problem. Our general results in many cases do not depend on the domain being binary or even finite. Similarly, it is possible to have finite or infinite constraint relations. This happens naturally, when we deal with algebraic constraints where constraints can be specified using algebraic (in)equations.

Example 2: Extensions of constraint satisfaction problems to quantified and stochastically quantified constraint satisfaction problems is done by allowing one to use first order quantifiers. Consider again the EO relation as defined in Example 1. An instance I of Q-SAT(EO) might look like $\forall x\ \exists y\ \forall z\ \exists w\ F$, where F is as defined above. Then it can be verified that I is not satisfiable. Moreover, $\mathbf{R}x\ \exists y\ \mathbf{R}z\ \exists w\ F$ is also not satisfiable. On the other hand $\forall x\ \exists y\ \forall z\ \exists w_1 \forall q \exists w_2\ ((EO(x,y,w_1) \wedge EO(z,y,w_2))$ is satisfiable.

Important Note: *The abstract contains only a discussion of the results, overview of techniques used and significance of the results. Full proofs and detailed definitions are developed in the full paper that can be obtained from the authors. Formal definitions of these problems can be found in [CF+93,CF+94,LMP99,HSM94,LMS96].*

2 Summary of results

As mentioned earlier, the focus of this paper is to develop a unified technique for characterizing the computational complexity and efficient approximability of quantified and stochastic satisfiability problems. For most part, we concentrate on the quantified and stochastic versions of the problems; the results for unquantified versions are derived in-situ. Specific results obtained in this paper are summarized in Figures 1 and 2. The general contributions of this paper include the following.

(1) An infinite class of quantified and stochastic constraint satisfaction problems is formalized. The type of problems studied include: decision, counting and optimization versions of these problems. Furthermore, combining these with the recent ideas of Littman et. al. we can define more general variants of the problem in which we vary the quantifiers and their semantics. We suspect that these infinite classes of problems will play a role similar to that already played by their unquantified counterparts in the earlier development of complexity theory. Of special note is the formalism of optimization and counting versions of these problems: these problems have not be been defined and studied in the literature prior to this paper. Recently there has been interest in studying the approximability of **PSPACE**-hard optimization problems:

(2) Two simple yet important concepts: *local replacements/reductions* and *relational representability* are formalized. We derive the basic complexity theoretic properties related to these concepts. Using these concepts, we propose unified methods for characterizing simultaneously, the decision, optimization, approximate optimization and counting complexity of quantified and stochastic constraint satisfaction problems.

(3) We derive very general sufficient conditions and generic reductions that simultaneously show that the decision and the approximate optimization problems are hard for their respective complexity classes. There has been a recent interest in studying the approximability of PSPACE-hard optimization problems. Our general results yield an infinite set of maximization versions of stochastic and quantified constrained satisfaction problems that are **PSPACE**-hard to approximate beyond a certain fixed constant and another infinite set that are **PSPACE**-hard to approximate for any n^ϵ, $\epsilon > 0$. Since the influential paper by Papadimitriou and Yannakakis on **MAX SNP**, there has been interest in finding logical/algebraic characterization of **NP**-hard optimization problems that are hard to approximate within different factors. The results for MAX-Q-SAT(S), MAX-S-SAT(S) MAX-NSF-Q-SAT(**S**) and MAX-NSF-S-SAT(**S**) provide similar algebraic characterizations of quantified and stochastic **PSPACE**-hard optimization problems.

We now discuss some of the specific results obtained in this paper and simultaneously contrast them with known results from the literature. These results are summarized in Figures 1 and 2. In order to allow for an easy comparison

between the results obtained here and the results obtained earlier by other researchers, we summarize both the results. Previous results are summarized in Figure 1 and new results are summarized in Figure 2. Moreover, previous results and our results are in 1-1 correspondence in terms of the numbering used. So for instance, 3(b) in *Part 1*, summarizes the earlier result on non-approximability of MAX-Q-3SAT, our result is given as 3(b) in *Part 2*.
Much of this discussion, but by no means all, is limited to finite sets D, since all *hardness* results given here are *tight* when D is finite. Almost all resulting reductions are *local*. Thus, they are $O(n \cdot logn)$ time-, linear size-, and $O(logn)$ space-bounded.

Part 1: Summary of related results applicable to this paper

(1) [Sc78,CES85,MS81]:The problems 3SAT and 3-COLORABLE GRAPH are **NQL**-complete. The problems EX-3SAT, EXACTLY1-EX3MONOSAT, NAE-EX3SAT, GOLD'S-MONOTONE-3SAT are $\leq^{bw}_{logn}$-complete for **NP**.

(2) [Sc78,Pa85]:The problems SAT($\mathbf{S}$) and SAT$_C$($\mathbf{S}$) are **NP**-complete and the problems Q-SAT($\mathbf{S}$) and Q-SAT$_C$($\mathbf{S}$) are **PSPACE**-complete, for all finite sets $\mathbf{S}$ of finite-arity Boolean relations such that $Rep_C(\mathbf{S})$=BOOLEAN-RELATIONS [a]. The problem S-3SAT is **PSPACE**-complete.

(3) (a) [ALM+98,PY91]: The problems MAX-3SAT and MAX-NAE-3SAT are **MAX SNP**-complete. Consequently, there exists $\epsilon > 0$ for which approximating these two problems within ϵ times optimum is **NP**-hard.
(b) [CF+93]: $\exists\ \epsilon > 0$ for which approximating the problem MAX-Q-3SAT within ϵ times optimum is **PSPACE**-hard.
(c) [CF+94]: $\exists\ \epsilon > 0$ for which approximating the problem MAX-S-3SAT within ϵ times optimum is **PSPACE**-hard.

(4) (a) [PR93]:The problem MAX-NSF-3SAT is **MAX** Π_1-complete. Consequently for all $\epsilon > 0$ approximating this problem within ϵ times optimum is **NP**-hard.
(b) [CF+93]:For all $\epsilon > 0$ approximating the problem MAX-Q-FORMULA-3SAT within ϵ times optimum is **PSPACE**-hard.

[a] This is the terminology used in [Sc78] to say that we can represent all finite arity-Boolean relations.

Fig. 1. Summary of results for constrained satisfaction problems obtained earlier.

Part 2: Summary of the results obtained in this paper

Let $k \geq 2$. Let $\mathbf{S}$ be a finite set of finite-arity relations on D_k, where k denotes the size of D, such that $Rep(\mathbf{S}) = D_k$-RELATIONS. [a] Then the following hold:

(1) The problems SAT($\mathbf{S}$) and SAT$_C$($\mathbf{S}$) are both **NQL**-complete and $\leq^{bw}_{logn}$-complete for **NP**.

(2) The problems Q-SAT($\mathbf{S}$), Q-SAT$_C$($\mathbf{S}$), are **PSPACE**-complete. Letting $k = 2$, the problem S-SAT($\mathbf{S}$) and S-SAT$_C$($\mathbf{S}$) are **PSPACE**-complete.

(3) Let $k \geq 2$. Let $\mathbf{S}$ be any finite set of finite-arity relations on D_k such that $Rep(\mathbf{S}) = D_k$-RELATIONS. Then, the following hold:

   (a) The problem MAX-SAT($\mathbf{S}$) is **MAX SNP**-complete. Consequently, there exists $\epsilon > 0$ such that approximating the problem within ϵ times optimum is **NP**-hard.
   (b) $\exists\ \epsilon > 0$ for which approximating the problems MAX-Q-SAT($\mathbf{S}$) within ϵ times optimum is **PSPACE**-hard.
   (c) Letting $k = 2$, $\exists\ \epsilon > 0$ for which approximating the problems MAX-S-SAT($\mathbf{S}$) within ϵ times optimum is **PSPACE**-hard.

(4) Let $\mathbf{S}$ and $\mathbf{T}$be finite sets of finite-arity relations on an arbitrary nonempty set D. Let $\epsilon > 0$. Then, the following hold:

   (a) The problem SAT($\mathbf{S}$) is $O(n \cdot logn)$ time-, linear size-, and $O(logn)$ space-bounded reducible to the problem of approximating the problem MAX-NSF-SAT($\mathbf{S}$) within a factor of ϵ times optimum. Therefore whenever the problem SAT($\mathbf{S}$) is **NP**-hard, approximating the problem MAX-NSF-SAT($\mathbf{S}$) within ϵ times optimum is **NP**-hard.
   (b) The problems Q-SAT($\mathbf{S}$), Q-SAT(S) are $O(n \cdot logn)$ time-, linear size-, and $O(logn)$ space-bounded reducible to the problems of approximating the problems MAX-NSF-Q-SAT($\mathbf{S}$), MAX-NSF-S-SAT($\mathbf{S}$), respectively, within a factor of n^{ϵ} times optimum.

[a] Like in Boolean case, this means that all finite arity relations on D_k can be equivalently represented as finite existentially-quantified conjunctions of relations in $\mathbf{S}$ applied to variables (to variables and constant symbols in C).

Fig. 2. Summary of our results for constrained satisfaction problems. The results are organized in a 1-1 correspondence with the results in Figure 1.

2.1 Extensions

As mentioned earlier, the results outlined here can be extended in at least two different directions. We briefly discuss these extensions.

Succinctly specified quantified formulas. By using the succinct specifications studied in [LW92,MH+94,Or82], it is possible to define hierarchically and periodically/dynamically specified quantified and stochastically quantified formulas. The techniques developed here and in [HSM01a,HSM01b,MH+94] can be combined to yield appropriate upper and lower bounds for such problems.

Quantified System of Equations. Let **F** be an algebraic structure. The formalism of constraint satisfaction problems can be extended to define the We denote the problem of determining the number of satisfying assignment of an **S**-formula, determining the solvability of a system of equations on F and counting the number of solutions to a system of equations on an algebraic structure F by #-SAT(**S**), SAT(**F**) and #-SAT(**F**) respectively. We define the corresponding quantified and stochastically quantified versions by #-Q-SAT(**S**), Q-SAT(**F**), #-Q-SAT(**F**), #-S-SAT(**S**), S-SAT(**F**) and #-S-SAT(**F**) respectively. The techniques developed here and in [HSM01a] can be naturally adapted to characterize the computational complexity of such problems. Based on the comprehensive discussion [LMP99,LMS96], we believe that these problems provide a rich collection of problems that might have wide applicability in AI literature.

2.2 Discussion and Significance

We discuss some of the above specific results in some detail. Note that some of the results that follow as corollaries of our general theorems have also been obtained previously by us or other researchers. Our purpose here is to demonstrate the effectiveness of the unified approach and to show that general results presented contain much of the earlier results as subsets of the general results. Moreover, the unified approach yields a large collection of new results that are reported for the first time in the literature. We make the following additional observations about the results summarized above.

First, note that several simple but fundamental properties of our model, that generalize those of previous models such as the *generalized CNF satisfiability problems*, the *constrained satisfiability problems*, and the classes of graphical problems *ECC* and *LCC* of [Sc78,FV93,JCG97,CJ+00,JCG97], respectively.

1. Most of our constructions hold, for domains D of arbitrary *not* necessarily finite cardinality. Moreover, they hold for problems expressed in terms of fairly arbitrary sets of algebraically-expressed constraints **S** on D. In particular, these sets of constraints also need not be finite.

2. Most of our constructions use the Boolean operator *and*, only in the sense of *simultaneously satisfiable* over the domain D and given set of constraints from **S**.

3. All of our constructions are explicitly expressed as *strongly-local graph /hypergraph replacements*. This guarantees their extensibility.

Second, the problems MAX-Q-SAT(S) and MAX-S-SAT(S) are **PSPACE**-hard (as opposed to **NP**-hard) to approximate beyond a fixed constant (a separate constant for each problem). Moreover MAX-NSF-Q-SAT(S) and MAX-NSF-S-SAT(S) are **PSPACE**-hard within any n^ϵ factor. Thus our results provide natural algebraic classes of optimization problems that can be potentially used for proving non-approximability of **PSPACE**-hard optimization problems. The un-quantified version of these problems have been used in the past to derive a number of non-approximability results. Similar results can be now obtained in a game theoretic setting.

Third, except for results in [FV93,JCG97] on when the problems SAT(**S**) are polynomially solvable and the well-known results that, the problems k-Colorable-Graph and MAX-k-Colorable-Graph are **NP**- and **MAX SNP**-complete, respectively, very few general hardness results were known previously for sets of relations on sets D such that $3 \leq |D| < \infty$.

Fourth, most of our reductions showing *hardness* are reductions that relate *simultaneously* the complexities of decision, counting, maximization, or approximate maximization variants of the variant constraint satisfaction problems. For e.g. the same basic reduction can be simultaneously used to show that Q-1-3SAT is **PSPACE**-hard, MAX-Q-1-3SAT is **PSPACE**-hard to approximate beyond a certain constant, #-Q-1-3SAT is #-**PSPACE**-hard, etc. Furthermore, generally, the reduction simultaneously also preserve the underlying graph theoretic structure of the problem instances. For want of a better term, we call these *single* multi-purpose reductions *SIMULTANEOUS-reductions.*

Our results extend earlier results and/or answer open problems in the literature. These include: (i) Ladner [La89,BMS81] to identify new natural **#PSPACE**-hard and -complete counting problems, (ii)Condon, Feigenbaum, Lund and Shor [CF+93,CF+94] to identify natural classes of **PSPACE**-hard optimization problems with provably **PSPACE**-hard ϵ-approximation problems, (iii) work of Papadimitriou [Pa85] on stochastic satisfiability problems (where only S-3SAT was considered) and (iv) Schaefer [Sc78] on quantified generalized satisfiability problems extending it to non-Boolean domain and providing tighter reductions). Progress is made on the approximability of the problems MAX-S-SAT(**S**) and MAX-Q-SAT(**S**): *a significant step towards obtaining a dichotomy theorems* for these problems. recently there has been substantial interest in obtaining dichotomy results for decision, optimization and counting versions of satisfiability problems. [CF+93,CF+94,Cr95,KSW97,LMP99]. While (non)-approximability of **NP**-hard optimization problems has received a lot of attention over the recent years, approximability of **PSPACE**-hard optimization problems has only been studied by us [HSM94,MH+94] for quantified and succinctly specified problems, by Condon, Feigenbaum, Lund and Shor [CF+93,CF+94] for quantified and stochastic satisfiability problems and by Lincoln, Mitchell and Scederov in the context of linear logic [LMS96].

3 Overall technique

Our methodology is based upon the following two *simple yet powerful* concepts.

1. Relational Representability: As the name suggests, letting $\mathbf{S}$ and $\mathbf{T}$ be sets of relations/algebraic constraints on a common domain D, the intuitive concept that the relations in $\mathbf{S}$ are *expressible* (or extending the terminology from [Sc78] are *representable*) by finite conjunctions of the relations in $\mathbf{T}$. This is formalized in Definition 1 below:

Definition 1 *(1) We denote the set of all finite-arity relations on a non-empty set D logically equivalent to finite existentially-quantified conjunctions of relations/algebraic constraints in* $\mathbf{S}$ *applied to variables (to variables and constant symbols in C) by* $Rep(\mathbf{S})$ *(by* $Rep_C(\mathbf{S})$*.)*

(2) We say that a relation S is, and a set of relations $\mathbf{S}$ *are, representable (constant-free representable) by a set of relations/algebraic constraints* $\mathbf{T}$ *if and only if* $S \in Rep_C(\mathbf{T})$ *(*$\mathbf{S} \in Rep(\mathbf{T})$*) and* $S \subset Rep_C(\mathbf{T})$ *(*$\mathbf{S} \subset Rep(\mathbf{T})$*, respectively.)*

Note: *Throughout this paper* $Rep(\mathbf{S})$ *denotes the set of relations expressible by* constant-free $\mathbf{S}$*-formulas; and* $Rep_C(\mathbf{S})$ *denotes the set of relations expressible by* $\mathbf{S}$*-formulas with constants from C.*

Variants of the concepts of Definition 1 on the *relative representability* of ordered-pairs $(\mathbf{S}, \mathbf{T})$ of sets of *relations*, henceforth denoted collectively by *relational representability*, are well known, especially in mathematical logic. Previously in complexity theory, *relational representability* as used here and the individual constraint satisfaction problems studied have usually been restricted to finite sets $\mathbf{S}$ of finite-arity relations on *finite* sets D, generally the set $\{0, 1\}$. Additionally, their uses are generally restricted to formulas or (occasionally also to quantified formulas), [Ho97,CES85,GJ79,JCG97,Sc78]. In contrast, our results apply with the exception of the problems S-SAT($\mathbf{S}$) to *both* finite and infinite domains and sets of relations/constraints.

For any set D and finite sets of finite-arity relations $\mathbf{S}$ and $\mathbf{T}$ on D, if $\mathbf{S} \subset Rep(\mathbf{T})$ (or $\mathbf{S} \subset Rep_C(\mathbf{T})$), then

(1) the problem SAT($\mathbf{S}$) is *1-strongly local* reducible to the problem SAT($\mathbf{T}$) (or $\mathrm{SAT}_C(\mathbf{T})$),

(2) the problem Q-SAT($\mathbf{S}$) is efficiently reducible to the problem Q-SAT($\mathbf{T}$) (or Q-$\mathrm{SAT}_C(\mathbf{T})$), and

(3) (when D is finite) the problem S-SAT($\mathbf{S}$) is efficiently reducible to the problem S-SAT($\mathbf{T}$) (or S-$\mathrm{SAT}_C(\mathbf{T})$).

(4) Moreover often, the reductions of items 1-3 can also be used to relate the relative complexities of the associated MAX- problems.

Fig. 3. *Meta-Result 2.* Relational Representability and Strongly-Local Reductions.

2. Local Replacements: Let $k \geq 1$. The second basic component of our methodology consists of the formalization and systematic investigation of the properties of the classes of k*-strongly-local* and k*-strongly-local-enforcer replacements* and *reductions*, especially with respect to constraint satisfaction problems. The basic idea of local reductions is not new and can be traced back to [GJ79] for decision problems, and recently in [HSM94,KSW97,Cr95] for optimization problems. The new contribution of this and companion papers is to formalize the complexity theory properties of such reductions. *In contrast, previous researchers, e.g. [GJ79,CES85], have discussed efficient reductions by local replacement; but they have* not *gone far in formalizing, or in characterizing the complexity-theoretic properties of, their concepts.*

Let $k \geq 1$. Let D_1,D_2 be nonempty sets. Let $\mathbf{S}$ with $|\mathbf{S}| = p$ and $\mathbf{T}$ with $|\mathbf{T}| = q$ be finite nonempty sets of finite-arity relations on D_1 and D_2, respectively. We define *k-strongly-local* and *k-strongly-local-enforcer reductions* of the problem SAT($\mathbf{S}$) to the problem SAT($\mathbf{T}$) to be *k-strongly-local* and *k-strongly-local-enforcer replacements* from the set of all $\mathbf{S}$-formulas to the set of all $\mathbf{T}$-formulas, that are also reductions Intuitively, $\forall k$, in *k-strongly-local replacements* we have *templates*, to be treated as *macros*, with the same template for each variable and distinct templates for each $\mathbf{S}$-clause. Details about macro expansions and the the way the variables are replaced depend very simply on the value of k. Figure 3 shows how local replacement/reductions and relational representability can be combined to obtain efficient reductions between classes of satisfiability problems.

Acknowledgement: We thank Anne Condon, Joan Feigenbaum and Gabriel Istrate for interesting discussions on related topics and pointers to related literature.

References

[AB88] D. Arnon and B. Buchberger(Editors), *Algorithms in Real Algebraic Geometry*, Academic Press, London, 1988.

[ALM+98] S. Arora, C. Lund, R. Motwani, M. Sudan and M. Szegedy. Proof Verification and Hardness of Approximation Problems. *J.of the ACM (JACM)*, 45, 1998, pp. 501-555.

[AC+98] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela and M. Protasi. *Complexity and Approximation: Combinatorial Optimization Problems and Their Approximability Properties*, Springer Verlag, 1999.

[BMS81] A. Bertoni, G. Mauri, and N. Sabadini. A characterization of the class of functions computable in polynomial time on random access machines. *Proc. 13th Annual ACM Sym. on the Theory of Computing*, 1981, pp. 168-176.

[Bo88] H. Bodlaender, Dynamic programming on graphs of bounded treewidth. *Proc. 15th International Colloquium on Automata Languages and Programming (ICALP)*, LNCS Vol. 317, 1988, pp. 105-118.

[BHR84] P.A. Bloniarz, H.B. Hunt III, and D.J. Rosenkrantz, Algebraic Structures with Hard Equivalence and Minimization Problems. *J. of the ACM (JACM)*, 31, 1984, pp. 879-904.

[CES85] M. Chung, W. Evangelist and I. Sudborough.

Complete problems for space bounded subclasses of NP. *Acta Informatica*, 22, 1985, pp. 379-395.

[CJ+00] D. Cohen, P. Jeavons, P. Jonsson, and M. Koubarakis Building Tractable Disjunctive Constraints *Journal of the ACM (JACM)*, 47 (2000), pp. 826-853.

[CF+93] A. Condon, J. Feigenbaum, C. Lund and P. Shor, Probabilistically Checkable Debate Systems and Approximation Algorithms for PSPACE-Hard Functions. *Chicago Journal of Theoretical Computer Science,* Vol. 1995, No. 4.

[CF+94] A. Condon, J. Feigenbaum, C. Lund and P. Shor. Random debaters and the hardness of approximating stochastic functions. *SIAM J. Computing*, 26, 1997, pp. 369-400.

[Cr95] N. Creignou. A dichotomy theorem for maximum generalized satisfiability problems. *J. of Computer and System Sciences (JCSS)*, 51, 1995, pp. 511-522.

[FV93] T. Feder and M. Vardi. Monotone monadic SNP and constraint satisfaction. *Proc. 25th Annual ACM Symposium on the Theory of Computing (STOC)*, 1993, pp. 612-622.

[FY79] A. Fraenkel and Y. Yesha. Complexity of problems in games, graphs, and algebraic equations. *Discrete Mathematics,* 1, 1979, pp. 15-30.

[GJ79] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, San Francisco, 1979.

[Ho97] D. Hochbaum(Ed.). *Approximation Algorithms for NP-Hard Problems,* PWS Publishing Company, Boston, MA, 1997.

[HSM94] H. Hunt III, R. Stearns and M. Marathe. Generalized CNF satisfiability problems and non-efficient approximability. *Proc. 9th Annual IEEE Conf. on Structure in Complexity Theory*, 1994, pp.356-366.

[HSM01a] H. Hunt III, R. Stearns and M. Marathe. Relational Representability, Local Reductions and the Complexity of Generalized Satisfiability Problem. submitted. Technical Report No. LA-UR-00-6108, Los Alamos National Laboratory.

[HSM01b] H. Hunt III, R. Stearns and M. Marathe. Strongly Local Reductions and the Complexity/Efficient Approximability of Algebra and Optimization on Abstract Algebraic Structures. *International Conference on Symbolic and Algebraic Computations (ISSAC),* July 2001.

[IM83] O. Ibarra and S. Moran. Probabilistic algorithms for deciding equivalence of straight-line programs. *J. of the ACM (JACM)* Vol.30, 1983, pp.217-228.

[JCG97] P. Jeavons, D. Cohen and M. Gyssens. Closure properties of constraints. *Journal of the ACM, (J. ACM)* 44(4):527-548, July 1997.

[Kr88] M. Krentel. The Complexity of optimization problems. *J. Computer and System Sciences (JCSS)* 36, pp. 490-509, 1988.

[KSW97] S. Khanna, M. Sudan, D. Williamson. A Complete classification of the approximability of maximization problems derived from Boolean constraint satisfaction. *Proc. 29th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 11-20, 1997.

[La89] R. Ladner. Polynomial space counting problems. *SIAM J. Computing*, 18(6):1087-1097, December 1989.

[LW92] T. Lengauer and K. Wagner. The correlation between the complexities of non-hierarchical and hierarchical versions of graph problems. *J. Computer and System Sciences (JCSS)*, Vol. 44, 1992, pp. 63-93.

[LW87] T. Lengauer and C. Weiner. Efficient solutions of hierarchical systems of linear equations. *Computing*, 39, 1987, pp. 111-132.

[Le83] H.W. Lensra, Jr. Integer programming with a fixed number of variables. *Math. Oper. Res.*, 8, 1983, pp. 538-548.

[Li82] D. Lichtenstein. Planar Formulae and their Uses", *SIAM J. Computing*, Vol 11, No. 2, May 1982 , pp. 329-343.

[LMS96] P.D. Lincoln, J.C. Mitchell and A. Scederov. Optimization Complexity of Linear Logic Proof Games. *Theoretical Computer Science*, 227 (1999) pp. 299-331.

[LMP99] M. Littman, S. Majercik and T. Pitassi. Stochastic Boolean Satisfiability. *Journal of Artificial Intelligence*, Kluwer Publications, 2000.

[MS81] B. Monien and I. Sudborough. Bounding the bandwidth of NP-complete problems," *Theoretical Computer Science*, 41(2-3):141-167, December 1985.

[MH+94] M. Marathe, H. Hunt III, R. Stearns and V. Radhakrishnan, Approximation Algorithms for PSPACE-Hard Hierarchically and Periodically Specified Problems. *SIAM J. Computing*, 27(5), pp. 1237-1261, Oct. 1998.

[Or82] J. Orlin. The Complexity of dynamic/periodic languages and optimization Problems. *Proc. 13th ACM Annual Symposium on Theory of Computing (STOC)*, 1978, pp. 218-227.

[Pa85] C. Papadimitriou. Games against nature. *J. Computer and System Sciences (JCSS)* , 31, 1985, pp. 288-301.

[PY91] C. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *J. Computer and System Sciences (JCSS)*, 43, 1991, pp. 425-440.

[Pa94] C. Papadimitriou. *Complexity Theory.* Addison-Wesley, Reading, MA, 1994.

[PR93] A. Panconesi and D. Ranjan. Quantifiers and approximations. *Theoretical Computer Science (TCS)*, 107, 1993, pp. 145-163.

[Sc78] T. Schaefer. The complexity of satisfiability problems. *Proc. 10th Annual ACM Symposium on Theory of Computing (STOC)*, 1978, pp. 216-226.

[Zu93] D. Zuckerman. On Unapproximable Versions of NP-Complete Problems. *SIAM J. Computing,* 25(6), pp. 1293-1304, 1996.