Executive Summary of Dagstuhl-Seminar

Complexity of Constraints, 06401

01/10/06-06/10/06

Organizers:	Nadia Creignou	U. Méditerranée Marseille, FR
	Phokion G. Kolaitis	IBM Almaden Research Center
		and U. of California, Santa Cruz, USA
	Heribert Vollmer	U. Hannover, DE

Overview of the Seminar's Main Topic

In a constraint satisfaction problem, the goal is to find an assignment of values to a given set of variables so that certain specified constraints are satisfied. Constraint satisfaction problems were introduced in the 1970s to model computational problems encountered in scene processing. It was quickly realized, however, that constraint satisfaction gives rise to a powerful general framework in which a wide variety of combinatorial problems can be expressed. As a matter of fact, it has been asserted that "Constraint satisfaction has a unitary theoretical model with myriad practical applications" (A. Mackworth, Foreword to *Constraint Processing* by Rina Dechter, 2003). Indeed, in artificial intelligence and other areas of computer science the constraint satisfaction framework is widely regarded as a versatile and efficient way of modeling and solving a number of real-world problems, such as planning and scheduling, frequency assignment problems, image processing, programming language analysis, and natural language understanding. In database theory, it has been shown that the key problem of conjunctive-query evaluation is actually equivalent to the constraint satisfaction problem. Furthermore, certain central problems in combinatorial optimization can be represented as constraint satisfaction problems. Thus, nowadays constraint satisfaction problems (CSPs) are ubiquitous in many different areas of computer science, from artificial intelligence and database systems to circuit design, network optimization, and theory of programming languages. Consequently, it is important to analyze and pinpoint the computational complexity of certain algorithmic tasks related to constraint satisfaction. These include determining if a CSP has a solution (and, if so, finding such a solution), counting the number of solutions of a CSP, enumerating all solutions of a CSP, and finding the biggest number of constraints that can be simultaneously satisfied, if a CSP is unsatisfiable. Complexity-theoretic results about these tasks may have direct impact on, for instance, the design and processing of database query languages, or strategies in data-mining, or the design and implementation of planners.

During the past two decades, an impressive array of diverse techniques from mathematical fields, such as propositional logic, model theory, Boolean function theory, universal algebra and combinatorics, have been used to analyze the computational complexity of algorithmic tasks related to CSPs. Although significant progress has been made on several fronts, some of the central questions remain unsolved so far; perhaps the most prominent of these is to obtain a complete classification of the complexity of CSPs over an arbitrary, but fixed, finite domain. One of the main aims of the Dagstuhl Seminar on *Complexity of Constraints* was to bring together researchers from all areas of activity in constraint satisfaction, so that they can communicate state-of-the-art advances and embark on a systematic interaction that will enhance the synergy between the different areas.

Detailed Description of the Seminar's Topics

An instance of a constraint satisfaction problem consists of a finite set $V = \{x_1, \ldots, x_n\}$ of variables, a finite domain U of possible values for the variables, and a finite set $\Phi = \{C_1, \ldots, C_m\}$ of constraints that restrict the combinations of values that certain tuples of variables may take. More precisely, each constraint is given by a relation R on the domain U applied to a tuple of r variables from U. We say that such an instance is *satisfiable* if there is an assignment I of elements from U to the variables in V such that for every $R(x_1, \ldots, x_r) \in \Phi$, the tuple $(I(x_1), \ldots, I(x_r))$ is in R. For example, let $U = \{1, \ldots, k\}$ be a k-element domain and let $\mathrm{NEQ} = \{(a, b) \mid a, b \in U, a \neq b\}$ be the complement of the equality relation on U. In this case, a typical set of constraints is $\Phi = \{\mathrm{NEQ}(x_1, x_2), \mathrm{NEQ}(x_1, x_4), \mathrm{NEQ}(x_2, x_4), \mathrm{NEQ}(x_2, x_3)\}$. It is easy to see that the GRAPH COLORABILITY can be expressed using constraints involving the relation NEQ; indeed, a graph $G = (\{1, \ldots, n\}, E)$ is k-colorable iff $\Phi = \{\mathrm{NEQ}(x_i, x_j) \mid (i, j) \in E\}$ is satisfiable.

It is known that the constraint satisfaction problem can also be formalized as the HOMOMORPHISM PROBLEM: given two finite relational structures A and S, is there a homomorphism h from A to S? Intuitively, the structure A consists of the variables and the constrained combinations of variables, while the structure S consists of the domain and the relations that provide the combinations of values that the constrained combinations of variables may take. A different, but equivalent, way to formalize CSP is to identify it with the CONJUNCTIVE QUERY EVALUATION PROBLEM: given a conjunctive query q and a finite relational structure S (that is, a relational database), does S satisfy q? Here, a *conjunctive query* is a first-order formula of the form $\exists x_1 \ldots \exists x_m \psi$, where ψ is a conjunction of atomic formulas with predicates from S. In other words, from a database point of view, the constraint relations are are just relational tables, and a CSP instance is nothing else than a *select-project-join* query in relational algebra. Thus, CSP is also a fundamental problem in database query processing, since conjunctive queries are the most frequently asked queries in databases.

In its full generality, CSP is an NP-complete problem, since, as seen above, it contains GRAPH COLORABILITY as a special case; for this reason, there has been an extensive pursuit of polynomial-time cases of CSP that are often called "islands of tractability" of constraint satisfaction. The input to a CSP problem consists of two parts: in the formulation of CSP as the HOMOMORPHISM PROBLEM, the two parts are a structure A and a structure S, while in the in the formulation of CSP as the CONJUNCTIVE QUERY EVALUATION PROBLEM the two parts are a conjunctive query q and a structure S. The discovery of islands of tractability of CSP entails finding restrictions on one or on both parts of the input that imply polynomial-time solvability. For example, it has been shown that CSP is solvable in polynomial-time if the input structure A is of bounded treewidth, while at the same time the structure S can be arbitrary. This is an important result in both artificial intelligence and database theory, as it yields some of the largest islands of tractability of CSP.

Note that the preceding discussion is about *uniform* CSP, in the sense that the input consists of both structures A and S (or both a conjunctive query q and a structure S). A *parametrized* family of *non-uniform* constraint satisfaction problems $\mathrm{CSP}(S)$ can be obtained by fixing the structure S, where the input to each of these problems is a structure A in the first formulation, or a conjunctive query q in the second formulation. Much of the research on the complexity of $\mathrm{CSP}(S)$ is motivated from the *dichotomy conjecture* of Feder and Vardi, which asserts that for each structure S, the decision problem $\mathrm{CSP}(S)$ is either NP-complete or solvable in polynomial time. In 1978, Schaefer proved a seminal Dichotomy Theorem for generalized satisfiability problems; this settles the Feder-Vardi dichotomy conjecture for $\mathrm{CSP}(S)$ when the universe of S is the Boolean domain $U = \{0, 1\}$.

In 2002, Bulatov established the dichotomy conjecture for structures S over a three element domain. This conjecture, however, remains open for structures with domains of cardinality bigger than three.

Schaefer's Dichotomy Theorem for the computational complexity of CSP(S) over the Boolean domain $U = \{0, 1\}$ has motivated the investigation of the computational complexity of other related algorithmic tasks over the Boolean universe. A number of interesting complexity results have been obtained concerning such computational tasks as counting the number of solutions, enumerating all solutions, computing solutions that are optimal in certain ways, and so on.

In addition to applying standard tools from computational complexity in this investigation, researchers have made a systematic use of a *generalized Galois correspondence* between certain sets S of Boolean constraint relations and closed classes of Boolean functions. Moreover, Post's lattice of all closed classes of Boolean functions has been fruitfully exploited more recently to completely classify the computational complexity of certain variants of constraint satisfaction. For example, in this context, the so called audit problem from database theory or the counting and enumeration problem have been classified.

The aforementioned Galois correspondence is also meaningful for finite domains other than the Boolean universe. For arbitrary finite domains U, function algebras and relational algebras over U have been considered, and deep results from *universal algebra* have been used to derive broad sufficient conditions for tractability of CSP(S); these conditions typically entail algebraic closure properties obeyed by the relations of the structure S.

The organizers felt that the seminar would provide a unique opportunity to focus attention on a number of important research problems in the complexity of constraints, including the following:

- *Islands of tractability of uniform CSP:* As mentioned earlier, the classes of structures of bounded treewidth constitute one of the largest islands of tractability of uniform CSP. Recently, these have been subsumed by some larger islands of tractability, such as the classes of structures of bounded hypertree-width. Much remains to be done, however, in discovering different or broader sufficient conditions for tractability of uniform CSP.

- *Complexity classifications for non-uniform CSP:* Despite much progress in the past years, many central complexity-theoretic questions for non-uniform CSP remain unsettled so far. The most prominent is the resolution of the Feder-Vardi dichotomy conjecture for CSP(S) over arbitrary, but fixed, finite domains. As mentioned earlier, this conjecture has been established for domains with at most three elements, but only partial results are known for larger domains. In particular, for many structures S, efficient algorithms for CSP(S) are known, while for some others, NP-completeness has been established, but a complete classification is still open. For other related algorithmic tasks such as counting, enumeration and optimization only a few results have been obtained thus far.

- *Quantified Constraint Satisfaction:* Recently, there has been growing activity in quantified constraint satisfaction where both existential and universal quantification over constrained variables is allowed, instead of just existential quantification. Equivalently, this is the extension of the CONJUNCTIVE QUERY EVALUATION PROBLEM to the MODEL CHECKING PROBLEM for positive disjunction-free first-order formulas. This area is another contact point of complexity, logic and universal algebra. While a complete classification of quantified constraint satisfaction over the Boolean domain has been obtained, much remains to be done for bigger domains.

– *Study of complexity classes through the lens of Boolean CSP:* Constraint satisfaction problems have attracted much attention in complexity theory for two reasons: first, various variants of CSPs lie at the heart of many standard complexity classes and often give rise to prototypical complete problems for these complexity classes; second, despite their great expressiveness, they tend to avoid "intermediate" complexity, that is, they tend to be either tractable or complete for standard complexity classes. In particular, this is true for Boolean CSPs. Thus, by focusing on this restricted world, it is possible to present a reasonably accurate bird's eye view of complexity theory and of the classes it has created.

Even in the restricted Boolean case, some algorithmic problems deserve further study. As examples, let us mention the satisfiability and the counting problem for extensions of propositional logic of interest in artificial intelligence and verification, such as circumscription and modal logics. Classification results here will be of interest in computational complexity theory, because they will shed new light on the structure of counting classes ($\#P$, $\# \cdot NP$) and on the power of reducibilities used here (many-one vs. subtractive vs. truth-table vs. Turing); moreover, they will be of interest in application areas, such as artificial intelligence, where these logics are of prime importance.

Organization of the Seminar and Activities

Most of the researchers invited to participate responded positively; at the end, the Seminar brought together forty one (41) researchers from different areas of constraint satisfaction and with complementary expertise. The list of participants consisted of both senior and junior researchers, including a number of postdoctoral researchers and a smaller number of advanced graduate students.

The organizers asked seven participants to present hour-long survey talks early in the week. The purpose of these talks was to cover the most important tools and techniques that are currently used in the study of CSP, as well as the different perspective (algebraic, logical, or graph-theoretic) that can be taken in investigating the complexity of constraints. The presenters and the topics of these survey talks are as follows:

1. H. Vollmer (U. Hannover, DE), Boolean CSP

2. F. Börner (U. Postdam, DE), Basic Universal Algebra and CSP

3. A. Bulatov (Simon Fraser U. , CA), Universal Algebra and Non-Uniform CSP

4. M. Valeriote (McMaster U., CA), Advanced Universal Algebra and CSP

5. A. Krokhin (U. Durham, UK), CSP and Dualities

6. F. Scarcello (U. della Calabria, IT), Uniform CSP and its Relationship to Database Theory

7. O. Kullmann (U. Swansea, UK), SAT Algorithms and Solvers

After this, the remaining participants were invited to present their recent work and to communicate state-of-the-art advances. There were twenty three (23) such communications, each ranging from 30 to 45 minutes. These communications focused on a number of important topics in the complexity of constraints:

- Algebraic methods and Galois correspondences

  -- I. Schnoor (U. Hannover, DE), New Algebraic Methods for CSPs
  -- F. Boerner (U. Postdam, DE), Different Galois Correspondences Relevant for CSPs

- Enumeration and counting

  -- H. Schnoor (U. Hannover, DE), Enumerating All Solutions of a CSP
  -- A. Durand (U. Paris 7, FR), Enumeration Complexity of Queryproblems and Quantifier Elimination
  -- P. Tesson (McGill U., CA), Systems of Equations over Semigroups and the #CSP-Conjecture

- The logical perspective and connections to database theory

  -- B. Larose (U. Concordia, CA), Non-FO CSP's are LOGSPACE-hard
  -- A. Dawar (U. Cambridge, UK), Definability of CSPs in Fixed-point Logics
  -- A. Atserias (UPC Barcelona, ES), On the Power of k-Consistency
  -- F. Madelaine (U. Durham, UK), Constraint Satisfaction, Structure Homomorphisms, and Fragments of Second Order Logic
  -- M. Gyssens (U. Limburg, BE), Survey on Hinges
  -- J. Niehren (INRIA Lille, FR), On XPath Dialects with Variables
  -- D. Itsykson (Steklov institute, RU), Exponential Lower Bound for Lovasz-Schrijver Proof Systems

- Optimization problems

  -- L. Kirousis (U. Patras, GR), Random Constraint Optimization: The case of Max-Cut
  -- G. Nordh (U. Linköping, SE), Complexity of the Maximal Solution Problem

- Complexity classifications and Dichotomy Theorems

  -- E. Maneva (U. Berkeley, US), Connectivity of Boolean Satisfiability: Structural and Computational Dichotomies
  -- S. Szeider (U. Durham, UK), Fixed-parameter Tractable Constraint Satisfaction
  -- H. Chen (U. Pompeu Fabra, ES), Quantified Equality Constraints
  -- M. Hermann (Ecole polytechnique, FR), Complexity of Clausal Constraints over Chains
  -- M. Bodirsky (Humboldt U Berlin, DE), Complexity of Temporal Constraint Satisfaction
  -- M. Grohe (Humboldt U Berlin, DE), Constraint Satisfaction with Succinctly Specified Relations
  -- D. Cohen (Royal Holloway, UK), The Expressibility Problem for Valued Constraints

- Algorithms for Satisfiability Problems

  -- E. Maneva (U. Berkeley, US), Survey Propagation
  -- O. Kullmann (U. Swansea, UK), CSPs in Clausal Form: Autarkies, Minimal Unsatisfiability, Hypergraphs

Concluding Remarks and Future Plans

The Dagstuhl Seminar on *Complexity of Constraints* took place a little more than six months after an International Workshop on *Mathematics of Constraint Satisfaction: Algebra, Logic and Graph Theory* organized by the University of Oxford, UK in March 2006
(see http://www.newton.cam.ac.uk/programmes/LAA/laaw03.html). There was a substantial overlap between the participants in this earlier workshop and those in the Dagstuhl Seminar on *Complexity of Constraints*. It turned out that bringing together researchers from different areas of theoretical computer science and mathematics initiated an interaction and that led to fruitful collaborations in attacking some of the open problems in this area. While much of the Oxford workshop consisted of survey talks and presentations of open problems, just six months later significant technical progress on several fronts was reported at the Dagstuhl seminar.

The organizers feel that that the Dagstuhl Seminar on *Complexity of Constraints* was a highly successful event at all levels. This is based on the plethora of new advances presented at Dagstuhl, the high quality of the presentations, and on very positive feedback about the seminar that the participents gave to the organizers at the end of the seminar and in subsequent email messages. In particular, the participants felt that the Dagstuhl Seminar was very stimulating and provided an impetus for continuing their efforts and interaction in advancing the state-of-the-art in this area. The organizers are planning to produce a Springer LNCS volume with survey papers based on the seven long survey talks and on some other talks that reported on more mature areas of research.

Finally, the organizers wish to sincerely thank the Scientific Directorate of the Center for its support of this event, and hope to have the opportunity to organize a follow-up seminar in the future.