Closure Properties of Constraints

PETER JEAVONS AND DAVID COHEN

Royal Holloway, University of London, London, United Kingdom

AND

MARC GYSSENS

University of Limburg, Diepenbeek, Belgium

Abstract. Many combinatorial search problems can be expressed as "constraint satisfaction problems" and this class of problems is known to be NP-complete in general. In this paper, we investigate the subclasses that arise from restricting the possible constraint types. We first show that any set of constraints that does not give rise to an NP-complete class of problems must satisfy a certain type of algebraic closure condition. We then investigate all the different possible forms of this algebraic closure property, and establish which of these are sufficient to ensure tractability. As examples, we show that all known classes of tractable constraints over finite domains can be characterized by such an algebraic closure property. Finally, we describe a simple computational procedure that can be used to determine the closure properties of a given set of constraints. This procedure involves solving a particular constraint satisfaction problem, which we call an "indicator problem."

Categories and Subject Descriptors: F.1.3 [**Computation by Abstract Devices**]: Complexity Classes—*reducibility and completeness*; F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Nonnumerical Algorithms and Problems—*complexity of proof procedures*; F.4.1 [**Mathematical Logic and Formal Languages**]: Mathematical Logic; G.2.1 [**Discrete Mathematics**]: Combinatorics

General Terms: Algorithms, Languages, Theory, Verification

Additional Key Words and Phrases: Complexity, constraint satisfaction problem, indicator problem, NP-completeness

---

Earlier versions of parts of this paper were published as JEAVONS, P., COHEN, D., AND GYSSENS, M. 1995. A unifying framework for tractable constraints. In *Proceedings of the 1st International Conference on Principles and Practice of Constraint Programming—CP '95* (Cassis, France, Sept.). Lecture Notes in Computer Science, vol. 976. Springer-Verlag, New York, pp. 276–291, and JEAVONS, P., COHEN, D., AND GYSSENS, M. 1996. A test for tractability. In *Proceedings of the 2nd International Conference on Constraint Programming—CP '96* (Boston, Mass., Aug.). Lecture Notes in Computer Science, vol. 1118. Springer-Verlag, New York, pp. 267–281.

This research was partially supported by EPSRC research grant GR/L09936 and by the British–Flemish Academic Research Collaboration Programme of the Belgian National Fund for Scientific Research and the British Council.

Authors' addresses: P. Jeavons and D. Cohen, Department of Computer Science, Royal Holloway, University of London, Egham, Surrey TW20 0EX, England; M. Gyssens, Department WNI, University of Limburg, B-3590 Diepenbeek, Belgium.

1. *Introduction*

Solving a constraint satisfaction problem is known to be NP-complete [Mackworth 1977]. However, many of the problems that arise in practice have special properties that allow them to be solved efficiently. The question of identifying restrictions to the general problem that are sufficient to ensure tractability is important from both a practical and a theoretical viewpoint, and has been extensively studied.

Such restrictions may either involve the *structure* of the constraints (in other words, which variables may be constrained by which other variables)[1] or they may involve the *nature* of the constraints (in other words, which combinations of values are allowed for variables that are mutually constrained).[2]

In this paper, we take the second approach, and investigate those classes of constraints that only give rise to tractable problems whatever way they are combined. A number of distinct classes of constraints with this property have previously been identified and shown to be maximal [Cooper et al. 1994; Jeavons et al. 1995; Jeavons and Cooper 1996].

In this paper, we establish that any class of constraints that does not give rise to NP-complete problems must satisfy a certain algebraic closure condition, and hence this algebraic property is a *necessary* condition for a class of constraints to be tractable (assuming that P is not equal to NP). We also show that many forms of this algebraic closure property are *sufficient* to ensure tractability.

As an example of the wide applicability of these results, we show that all known examples of tractable constraint classes over finite domains can be characterized by an algebraic condition of this kind, even though some of them were originally defined in very different ways.

Finally, we describe a simple computational procedure to determine the algebraic closure properties of a given set of constraints. The test involves calculating the solutions to a fixed constraint satisfaction problem involving constraints from the given set.

The work described in this paper represents a generalization of earlier results concerning tractable subproblems of the GENERALIZED SATISFIABILITY problem. Schaefer [1978] identified all possible tractable classes of constraints for this problem, which corresponds to the special case of the constraint satisfaction problem in which the variables are Boolean. The tractable classes described in Schaefer [1978] are special cases of the tractable classes of general constraints described below, and they are given as examples.

A number of tractable constraint classes have also been identified by Feder and Vardi [1993]. They define a notion of *width* for constraint satisfaction problems in terms of the logic programming language Datalog, and show that problems with bounded width are solvable in polynomial time. It is stated in Feder and Vardi [1993] that the problem of determining whether a fixed collection of constraints gives rise to problems of bounded width is not known to be decidable. However, it is shown that a more restricted property, called *bounded strict width*, is decidable, and in fact corresponds to an algebraic closure

[1] See, for example, Dechter and Pearl [1988], Freuder [1985], Gyssens et al. [1994], Montanari [1974], and Montanari and Rossi [1991].
[2] See, for example, Cooper et al. [1994], Jeavons et al. [1995], Jeavons and Cooper [1996], Feder and Vardi [1993], Kirousis [1993], Montanari [1974], van Beek [1992], and Van Hentenryck et al. [1992].

property of the form described here. Other tractable constraint classes are also shown to be characterized by a closure property of this type. This paper builds on the work of Feder and Vardi [1993] by examining the general question of the link between algebraic closure properties and tractability of constraints, and establishing necessary and sufficient conditions for these closure properties.

A different approach to identifying tractable constraints is taken in van Beek [1992], where it is shown that a property of constraints, referred to as 'row-convexity', together with path-consistency, is sufficient to ensure tractability in binary constraint satisfaction problems. It should be noted, however, that because of the additional requirement for path-consistency, row-convex constraints do *not* constitute a tractable class in the sense defined in this paper. In fact, the class of problems that contain only row-convex constraints is NP-complete [Cooper et al. 1994].

The paper is organized as follows: In Section 2, we give the basic definitions, and in Section 3 we define what we mean by an algebraic closure property for a set of relations and examine the possible forms of such a closure property. In Section 4, we identify which of these forms are *necessary* conditions for tractability, and in Section 5, we identify which of them are *sufficient* for tractability. In Section 6, we describe a computational method to determine the closure properties satisfied by a set of relations. Finally, we summarize the results presented and draw some conclusions.

2. *Definitions*

2.1. The Constraint Satisfaction Problem

Notation 2.1.1. *For any set D, and any natural number n, we denote the set of all n-tuples of elements of D by D^n. For any tuple $t \in D^n$, and any i in the range* 1 *to n, we denote the value in the ith coordinate position of t by $t[i]$. The tuple t may then be written in the form $\langle t[1], t[2], \ldots, t[n]\rangle$.*

A subset of D^n is called an *n-ary relation* over D.

We now define the (finite) constraint satisfaction problem that has been widely studied in the Artificial Intelligence community [Ladkin and Maddux 1994; Mackworth 1977; Montanari 1974].

Definition 2.1.2. An instance of a *constraint satisfaction problem* consists of

—a finite set of variables, V;
—a finite domain of values, D;
—a set of constraints $\{C_1, C_2, \ldots, C_q\}$.

Each constraint C_i is a pair (S_i, R_i), where S_i is a list of variables of length m_i, called the *constraint scope*, and R_i is an m_i-ary relation over D, called the *constraint relation*. (The tuples of R_i indicate the allowed combinations of simultaneous values for the variables in S_i.)

The length of the tuples in the constraint relation of a given constraint will be called the *arity* of that constraint. In particular, unary constraints specify the allowed values for a single variable, and binary constraints specify the allowed combinations of values for a pair of variables. A *solution* to a constraint satisfaction problem is a function from the variables to the domain such that the image of each constraint scope is an element of the corresponding constraint relation.

Deciding whether or not a given problem instance has a solution is NP-complete in general [Mackworth 1977] even when the constraints are restricted to binary constraints. In this paper we consider how restricting the allowed constraint relations to some fixed subset of all the possible relations affects the complexity of this decision problem. We therefore make the following definition:

Definition 2.1.3 For any set of relations, Γ, **CSP**(Γ) is defined to be the decision problem with

INSTANCE: An instance, P, of a constraint satisfaction problem, in which all constraint relations are elements of Γ.

QUESTION: Does P have a solution?

If there exists an algorithm that solves every problem instance in CSP(Γ) in polynomial time, then we shall say that CSP(Γ) is a *tractable* problem, and Γ is a *tractable* set of relations.

Example 2.1.4. The binary disequality relation over a set D, denoted $\neq_D$, is defined as

$$\neq_D = \{\langle d_1, d_2\rangle \in D^2 | d_1 \neq d_2\}.$$

Note that CSP($\{\neq_D\}$) corresponds precisely to the GRAPH $|D|$-COLORABILITY problem [Garey and Johnson 1979]. This problem is tractable when $|D| \leq 2$ and NP-complete when $|D| \geq 3$.

Example 2.1.5. Consider the ternary relation δ over the set $D = \{0, 1\}$ which is defined by

$$\delta = \{\langle 0, 0, 1\rangle, \langle 0, 1, 0\rangle, \langle 1, 0, 0\rangle, \langle 0, 1, 1\rangle, \langle 1, 0, 1\rangle, \langle 1, 1, 0\rangle\}.$$

The problem CSP($\{\delta\}$) corresponds precisely to the NOT-ALL-EQUAL SATISFIABILITY problem [Garey and Johnson 1979], which is NP-complete [Schaefer 1978].

Example 2.1.6. We now describe three relations that will be used as examples of constraint relations throughout the paper.

Each of these relations is a set of tuples of elements from the domain $D = \{0, 1, 2\}$, as defined below:

$$R_1 = \{ \begin{array}{l}\langle 0, 0\rangle,\\ \langle 1, 2\rangle,\\ \langle 0, 1\rangle,\\ \langle 2, 1\rangle\end{array} \qquad R_2 = \{ \begin{array}{l}\langle 0, 1, 2\rangle,\\ \langle 1, 2, 0\rangle,\\ \langle 2, 0, 1\rangle \quad \}\end{array}$$

$$R_3 = \{ \quad \langle 2, 2\rangle, \\ \langle 2, 1\rangle, \\ \langle 2, 0\rangle, \\ \langle 1, 2\rangle, \\ \langle 1, 1\rangle, \\ \langle 1, 0\rangle, \\ \langle 0, 2\rangle, \\ \langle 0, 1\rangle \quad \}$$

The problem CSP($\{R_1, R_2, R_3\}$) consists of all constraint satisfaction problem instances in which the constraint relations are all chosen from the set $\{R_1, R_2, R_3\}$.

The complexity of CSP(Γ) for arbitrary subsets Γ of $\{R_1, R_2, R_3\}$ will be determined using the techniques developed later in this paper (see Example 6.6).

2.2. OPERATIONS ON RELATIONS. In Section 4, we shall examine conditions on a set of relations Γ that allow known NP-complete problems to be reduced to CSP(Γ). The reductions will be described using standard operations from relational algebra [Codd 1970], which are described in this section.

Definition 2.2.1. We define the following operations on relations.

—Let R be an n-ary relation over a domain D and let S be an m-ary relation over D. The *Cartesian product* $R \times S$ is defined to be the $(n + m)$-ary relation

$$R \times S = \{\langle t[1], t[2], \ldots, t[n+m]\rangle | (\langle t[1], t[2], \ldots, t[n]\rangle \in R) \ \wedge \\ (\langle t[n+1], t[n+2], \ldots, t[n+m]\rangle \in S)\}.$$

—Let R be an n-ary relation over a domain D. Let $1 \leq i, j \leq n$. The *equality selection* $\sigma_{i=j}(R)$ is defined to be the n-ary relation

$$\sigma_{i=j}(R) = \{t \in R | t[i] = t[j]\}.$$

—Let R be an n-ary relation over a domain D. Let $i_1, \ldots, i_m$ be a subsequence of $1, \ldots, n$. The projection $\pi_{i_1, \ldots, i_m}(R)$ is defined to be the m-ary relation

$$\pi_{i_1, \ldots, i_m}(R) = \{\langle t[i_1], \ldots, t[i_m]\rangle | t \in R\}.$$

It is well known that the combined effect of two constraints in a constraint satisfaction problem can be obtained by performing a relational *join* operation [Codd 1970] on the two constraints [Gyssens et al. 1994]. The next result is a simple consequence of the definition of the relational join operation.

LEMMA 2.2.2. *Any relational join of relations R and S can be calculated by performing a sequence of Cartesian product, equality selection, and projection operations on R and S.*

In view of this result, it will be convenient to use the following notation.

NOTATION 2.2.3. *The set of all relations that can be obtained from a given set of relations*, Γ, *using some sequence of Cartesian product*, *equality selection*, *and projection operations will be denoted* Γ^+.

Note that Γ^+ contains exactly those relations that can be obtained as 'derived' relations in a constraint satisfaction problem instance with constraint relations chosen from Γ [Cohen et al. 1996].

3. *Closure Operations*

We shall establish below that significant information about CSP(Γ) can be determined from algebraic properties of the set of relations Γ. In order to describe these algebraic properties we need to consider arbitrary operations on D, in other words, arbitrary functions from D^k to D, for arbitrary values of k.

For the results below, we shall be particularly interested in certain special kinds of operations. We therefore make the following definition:

Definition 3.1. Let $\otimes$ be a k-ary operation from D^k to D.

—If $\otimes$ is such that, for all $d \in D$, $\otimes(d, d, \ldots, d) = d$, then $\otimes$ is said to be *idempotent.*

—If there exists an index $i \in \{1, 2, \ldots, k\}$ such that for all $\langle d_1, d_2, \ldots, d_n \rangle \in D^k$ we have $\otimes(d_1, d_2, \ldots, d_n) = f(d_i)$, where f is a nonconstant unary operation on D, then $\otimes$ is called *essentially unary*. (Note that f is required to be nonconstant, so constant operations are *not* essentially unary.)
If f is the identity operation, then $\otimes$ is called a *projection.*

—If $k \geq 3$ and there exists an index $i \in \{1, 2, \ldots, k\}$ such that for all $d_1, d_2, \ldots, d_k \in D$ with $|\{d_1, d_2, \ldots, d_k\}| < k$ we have $\otimes(d_1, d_2, \ldots, d_k) = d_i$, but $\otimes$ is not a projection, then $\otimes$ is called a *semiprojection* [Rosenberg 1986; Szendrei 1986].

—If $k = 3$ and for all $d, d' \in D$ we have $\otimes(d, d, d') = \otimes(d, d', d) = \otimes(d', d, d) = d$, then $\otimes$ is called a *majority operation.*

—If $k = 3$ and for all $d_1, d_2, d_3 \in D$ we have $\otimes(d_1, d_2, d_3) = d_1 - d_2 + d_3$, where $+$ and $-$ are binary operations on D such that $\langle D, +, - \rangle$ is an Abelian group [Szendrei 1986], then $\otimes$ is called an *affine operation.*

Any operation on D can be extended to an operation on tuples over D by applying the operation in each coordinate position separately (i.e., pointwise). Hence, any operation defined on the domain of a relation can be used to define an operation on the tuples in that relation, as follows:

Definition 3.2. Let $\otimes$: $D^k \to D$ be a k-ary operation on D and let R be an n-ary relation over D.

For any collection of k tuples, $t_1, t_2, \ldots, t_k \in R$, (not necessarily all distinct) the n-tuple $\otimes(t_1, t_2, \ldots, t_k)$ is defined as follows:

$$\otimes(t_1, t_2, \ldots, t_k) = \langle \otimes(t_1[1], t_2[1], \ldots, t_k[1]),\ \otimes(t_1[2], t_2[2], \ldots, t_k[2]), \ldots, \otimes(t_1[n], t_2[n], \ldots, t_k[n]) \rangle.$$

Finally, we define $\otimes(R)$ to be the n-ary relation

$$\{ \otimes (t_1, \ldots , t_k) | t_1, \ldots , t_k \in R\}.$$

Using this definition, we now define the following closure property of relations.

Definition 3.3. Let $\otimes$ be a k-ary operation on D, and let R be an n-ary relation over D. The relation R is *closed* under $\otimes$ if $\otimes(R) \subseteq R$.

Example 3.4. Let Δ be the ternary majority operation defined as follows:

$$\Delta(x, y, z) = \begin{cases} y & \text{if} \quad y = z; \\ x & \text{otherwise.} \end{cases}$$

The relation R_2 defined in Example 2.1.6 is closed under Δ, since applying the Δ operation to any three elements of R_2 yields an element of R_2. For example,

$$\Delta(\langle 0, 1, 2\rangle, \langle 1, 2, 0\rangle, \langle 1, 2, 0\rangle) = \langle 1, 2, 0\rangle \in R_2.$$

The relation R_3 defined in Example 2.1.6 is *not* closed under Δ, since applying the Δ operation to the last three elements of R_3 yields a tuple which is not an element of R_3:

$$\Delta(\langle 1, 0\rangle, \langle 0, 2\rangle, \langle 0, 1\rangle) = \langle 0, 0\rangle \notin R_3. \qquad \square$$

For any set of relations Γ, and any operation $\otimes$, if every $R \in \Gamma$ is closed under $\otimes$, then we shall say that Γ is closed under $\otimes$. The next lemma indicates that the property of being closed under some operation is preserved by all possible projection, equality selection, and product operations on relations, as defined in Section 2.2.

LEMMA 3.5. *For any set of relations Γ, and any operation $\otimes$, if Γ is closed under $\otimes$, then Γ^+ is closed under $\otimes$.*

PROOF. Follows immediately from the definitions. $\square$

NOTATION 3.6. *For any set of relations Γ with domain D, the set of all operations on D under which Γ is closed will be denoted $\Gamma^{\triangleright}$.*

The set of closure operations, $\Gamma^{\triangleright}$, can be used to obtain a great deal of information about the problem CSP(Γ), as we shall demonstrate in the next two sections.

As a first example of this, we shall show that the operations in $\Gamma^{\triangleright}$ can be used obtain a reduction from one problem to another.

PROPOSITION 3.7. *For any set of finite relations Γ, and any $\otimes \in \Gamma^{\triangleright}$, there is a polynomial-time reduction from CSP(Γ) to CSP($\otimes(\Gamma)$), where $\otimes(\Gamma) = \{\otimes(R) | R \in \Gamma\}$.*

PROOF. Let $\mathcal{P}$ be any problem instance in CSP(Γ) and consider the instance $\mathcal{P}'$ obtained by replacing each constraint relation R_i of $\mathcal{P}$ by the relation $\otimes(R_i)$. It is clear that $\mathcal{P}'$ can be obtained from $\mathcal{P}$ in polynomial-time. It follows from Definition 3.3 that $\mathcal{P}'$ has a solution if and only if $\mathcal{P}$ has a solution. $\square$

It follows from this result that if $\Gamma^{\triangleright}$ contains a non-injective unary operation, then CSP(Γ) can be reduced to a problem over a smaller domain. One way to view this is that the presence of a noninjective unary operation in $\Gamma^{\triangleright}$ indicates that constraints with relations chosen from Γ allow a form of global "substitutability," similar to the notion defined by Freuder [1991].

If $\Gamma^{\triangleright}$ does *not* contain any noninjective unary operations, then we shall say that Γ is *reduced*. The next theorem uses a general result from universal algebra [Rosenberg 1986; Szendrei 1986] to show that the possible choices for $\Gamma^{\triangleright}$ are quite limited.

THEOREM 3.8. *For any reduced set of relations* Γ, *over a finite set*, *either*

(1) $\Gamma^{\triangleright}$ *contains essentially unary operations only*, *or*
(2) $\Gamma^{\triangleright}$ *contains an operation which is*

  (*a*) *a constant operation*; *or*
  (*b*) *a majority operation*; *or*
  (*c*) *an idempotent binary operation* (*which is not a projection*); *or*
  (*d*) *an affine operation*; *or*
  (*e*) *a semiprojection.*

PROOF. The set of operations $\Gamma^{\triangleright}$ contains all projections and is closed under composition; hence, it constitutes a "clone" [Cohn 1965; Szendrei 1986]. It was shown in Rosenberg [1986] that any nontrivial clone on a finite set must contain a minimal clone, and that any minimal clone contains either

(1) a nonidentity unary operation; or
(2) a constant operation; or
(3) a majority operation; or
(4) an idempotent binary operation (which is not a projection); or
(5) an affine operation; or
(6) a semiprojection.

Furthermore, if Γ is reduced, and $\Gamma^{\triangleright}$ contains *any* operations that are *not* essentially unary, then it is straightforward to show, by considering such an operation of the smallest possible arity, that $\Gamma^{\triangleright}$ contains an operation in one of the last five of these classes [Szendrei 1986; Machida and Rosenberg 1984]. □

In the next two sections, we examine each of these possibilities in turn, in order to establish what can be said about the complexity of CSP(Γ) in the various cases.

4. *A Necessary Condition for Tractability*

In this section, we will show that any set of relations that is only closed under essentially unary operations will give rise to a class of constraint satisfaction problems that is NP-complete.

THEOREM 4.1. *For any finite set of relations*, Γ, *over a finite set* D, *if* $\Gamma^{\triangleright}$ *contains essentially unary operations only then* CSP(Γ) *is NP-complete*.

PROOF. When $|D| \leq 2$, then we may assume without loss of generality that $D \subseteq \{0, 1\}$, where 0 corresponds to the Boolean value **false** and 1 corresponds to the Boolean value **true**. It follows that the problem CSP(Γ) corresponds to the GENERALIZED SATISFIABILITY problem over the set of Boolean relations Γ, as defined in Schaefer [1978] (see also Garey and Johnson [1979]).

It was established in Schaefer [1978] that this problem is NP-complete unless one of the following conditions holds:

(1) Every relation in Γ contains the tuple $(0, 0, \ldots, 0)$;
(2) Every relation in Γ contains the tuple $(1, 1, \ldots, 1)$;
(3) Every relation in Γ is definable by a formula in conjunctive normal form in which each conjunct has at most one negated variable;
(4) Every relation in Γ is definable by a formula in conjunctive normal form in which each conjunct has at most one unnegated variable;
(5) Every relation in Γ is definable by a formula in conjunctive normal form in which each conjunct contains at most 2 literals;
(6) Every relation in Γ is the set of solutions of a system of linear equations over the finite field GF(2).

It is straightforward to show that in each of these cases Γ is closed under some operation which is not essentially unary (see Jeavons [1995] for details). Hence, the result holds when $|D| = 2$.

For larger values of $|D|$, we proceed by induction. Assume that $|D| \geq 3$ and the result holds for all smaller values of $|D|$. Let $m = |D|(|D| - 1)$ and let $n = |D|^m$. Let M be an m by n matrix over D in which the columns consist of all possible m-tuples over D (in some order). Let R_0 be the relation consisting of all the tuples occurring as rows of M. The only condition we place on the choice of order for the columns of M is that $\pi_{1,2}(R_0) = \neq_D$, where $\neq_D$ is the binary disequality relation over D, as defined in Example 2.1.4.

We now construct a relation $\hat{R}_0$ which is the "closest approximation" to R_0 that we can obtain from the relations in Γ and the domain D using the Cartesian product, equality selection and projection operations:

$$\hat{R}_0 = \bigcap \{R \in (\Gamma \cup D^1)^+ | R_0 \subseteq R\}.$$

Since this is a finite intersection, and intersection is a special case of join, we have from Lemma 2.1.8 that $\hat{R}_0 \in (\Gamma \cup D^1)^+$. In other words, the relation $\hat{R}_0$ can be obtained as a derived constraint relation in some problem belonging to CSP(Γ).

There are now two cases to consider:

(1) If there exists some tuple $t_0 \in \hat{R}_0$ with $t_0[1] = t_0[2]$, then we will construct, using t_0, an appropriate operation under which Γ is closed.

Define the function $\otimes$: $D^m \to D$ by setting $\otimes(d_1, d_2, \ldots, d_m) = t_0[j]$ where j is the unique column of M corresponding to the m-tuple $\langle d_1, d_2, \ldots, d_m \rangle$. We will show that Γ is closed under $\otimes$.

Choose any $R \in \Gamma$, and let p be the arity of R. We are required to show that R is closed under $\otimes$. Consider any sequence $t_1, t_2, \ldots, t_m$ of tuples of R (not necessarily distinct), and, for $i = 1, 2, \ldots, p$, let c_i be the m-tuple

$\langle t_1[i], t_2[i], \ldots, t_m[i]\rangle$. For each pair of indices, i, j, such that $c_i = c_j$, apply the equality selection $\sigma_{i=j}$ to R, to obtain a new relation R'.

Now choose a maximal set of indices, $I = \{i_1, i_2, \ldots, i_s\}$, such that the corresponding c_i are all distinct, and construct the relation $R'' = \pi_I(R') \times D^{n-|I|}$. Finally, permute the coordinate positions of R'' (by a sequence of Cartesian product, equality selection, and projection operations), such that $R'' \supseteq R_0$ (this is always possible, by the construction of R_0 and R''). Since $R'' \in (\Gamma \cup D^1)^+$, we know that t_0 is a tuple of R'', by the definition of R_0. Hence the appropriate projection of t_0 is an element of R, and R is closed under $\otimes$.

If $\otimes$ is not essentially unary, then we have the result. Otherwise, let f: $D \to D$ be the corresponding unary operation, and set

$$f(D) = \{f(d) | d \in D\};$$

$$f(\Gamma) = \{\{\langle f(d_1), f(d_2), \ldots, f(d_r)\rangle | \langle d_1, d_2, \ldots, d_r\rangle \in C\} | C \in \Gamma\}.$$

By the choice of t_0, f cannot be injective, so $|f(D)| < |D|$. By the inductive hypothesis, we know that either CSP($f(\Gamma)$) is NP-complete (in which case CSP(Γ) must also be NP-complete) or else $f(\Gamma)$ is closed under some operation $\oplus$ that is not essentially unary (in which case Γ is closed under the operation $\oplus f$, which is also not essentially unary). Hence, the result follows by induction in this case.

(2) Alternatively, if $\hat{R}_0$ contains no tuple t such that $t[1] = t[2]$, then $\pi_{1,2}(\hat{R}_0) = \neq_D$, so $\neq_D \in (\Gamma \cup D^1)^+$. But this implies that CSP($\{\neq_D\}$) is reducible to CSP(Γ), since every occurence of the constraint relation $\neq_D$ can be replaced with an equivalent collection of constraints with relations chosen from Γ. However, it was pointed out in Example 2.1.4 that CSP($\{\neq_D\}$) corresponds to the GRAPH $|D|$-COLORABILITY problem [Garey and Johnson 1979], which is NP-complete when $|D| \geq 3$. Hence, this implies that CSP(Γ) is NP-complete, and the result holds in this case also. □

Combining Theorem 4.1 with Theorem 3.8 gives the following necessary condition for tractability.

COROLLARY 4.2. *Assuming that P is not equal to NP, any tractable set of reduced relations must be closed under either a constant operation, or a majority operation, or an idempotent binary operation, or an affine operation, or a semiprojection.*

Note that the arity of a semiprojection is at most $|D|$, so for any finite set D there are only finitely many operations matching the given criteria, which means that there is a finite procedure to check whether this necessary condition is satisfied (see Corollary 6.5).

5. *Sufficient Conditions for Tractability*

We have shown in the previous section that when Γ is a tractable set of relations, then $\Gamma^{\triangleright}$ must contain an operation from a limited range of types. For some of these types of operations, it can be shown that the relations in Γ have a particular structural form, which can be exploited to obtain an efficient algorithm for CSP(Γ).

We now consider each of the possibilities identified in Corollary 4.2 in turn, to determine whether or not they are sufficient to ensure tractability.

5.1. CONSTANT OPERATIONS. Closure under a constant operation is easily shown to be a sufficient condition for tractability.

PROPOSITION 5.1.1. *For any set of relations* Γ, *if* Γ *is closed under a constant operation*, *then CSP*(Γ) *is solvable in polynomial time*.

PROOF. If every relation in Γ is closed under some constant operation $\otimes$, with constant value d, then every nonempty relation in Γ must contain the tuple $\langle d, d, \ldots, d\rangle$. Hence, in this case, the decision problem for any constraint satisfaction problem instance in $\mathcal{P}$ in CSP(Γ) is clearly trivial to solve, since $\mathcal{P}$ either contains an empty constraint, in which case it does not have a solution, or else $\mathcal{P}$ allows the solution in which every variable is assigned the value d. $\square$

The class of sets of relations closed under some constant operation is a rather trivial tractable class. It is referred to in Jeavons et al. [1995] as *Class 0.*

Example 5.1.2. Let $\top$ denote the unary operation on the domain $D = \{0, 1, 2\}$ that returns the constant value 1. The constraint R_3 defined in Example 2.1.6 is closed under $\top$, since applying the $\top$ operation to any element of R_3 yields the tuple $\langle 1, 1\rangle$, which is an element of R_3. The constraint R_2 defined in Example 2.1.6 is *not* closed under $\top$, since applying the $\top$ operation to any element of R_2 yields the tuple $\langle 1, 1, 1\rangle$, which is not an element of R_2. In fact, R_2 is clearly not closed under any constant operation.

Example 5.1.3. When $D = \{\mathbf{true},\mathbf{false}\}$, there are only two possible constant operations on D.

The first two tractable subproblems of the GENERALIZED SATISFIABILITY problem identified by Schaefer [1978] correspond to the tractable classes of relations characterized by closure under these two constant operations.

5.2. MAJORITY OPERATIONS. We will now show that closure under a majority operation is also a sufficient condition for tractability.

We first establish that when a relation R is closed under a majority operation, any constraint involving R can be decomposed into binary constraints.

PROPOSITION 5.2.1. *Let R be a relation of arity n that is closed under a majority operation*, *and let C be any constraint* $C = (S, R)$ *with scope S and relation R.*

For any problem $\mathcal{P}$ *containing the constraint C*, *the problem* $\mathcal{P}'$ *obtained by replacing C with the set of binary constraints*

$$\{((S[i], S[j]),\ \pi_{i,j}(R)) \mid 1 \le i \le j \le n\}$$

has exactly the same solutions as $\mathcal{P}$.

PROOF. It is clear that any solution to $\mathcal{P}$ is a solution to $\mathcal{P}'$, since $\mathcal{P}'$ is obtained by taking binary projections of a constraint from $\mathcal{P}$.

Now let σ be any solution to $\mathcal{P}'$, and set $t = \langle \sigma(S[1]), \sigma(S[2]), \ldots, \sigma(S[n]) \rangle$. We shall prove, by induction on n, that $t \in R$, thereby establishing that σ is a solution to $\mathcal{P}$.

For $n < 3$ the result holds trivially, so assume that $n \geq 3$, and that the result holds for all smaller values. Let $I = \{1, 2, \ldots, n\}$ be the set of indices of positions in S and choose $i_1, i_2, i_3 \in I$. By Proposition 3.5 and the inductive hypothesis, applied to $\pi_{I \setminus \{i_j\}}(R)$, there is some $t_j \in R$ which agrees with t at all positions except i_j, for $j = 1, 2, 3$. Since R is closed under a majority operation, applying this operation to t_1, t_2, t_3 gives $t \in R$. $\square$

Example 5.2.2. Recall the relation R_2 defined in Example 2.1.6.

It was shown in Example 3.4 that R_2 is closed under the operation Δ. Since this operation is a majority operation, we know by Proposition 5.2.1 that any constraint with relation R_2 can be decomposed into an equivalent collection of binary constraints with the following relations:

$$\pi_{1,2}(R_2) = \{\langle 0, 1\rangle, \langle 1, 2\rangle, \langle 2, 0\rangle\}$$

$$\pi_{1,3}(R_2) = \{\langle 1, 0\rangle, \langle 2, 1\rangle, \langle 0, 2\rangle\}$$

$$\pi_{2,3}(R_2) = \{\langle 0, 1\rangle, \langle 1, 2\rangle, \langle 2, 0\rangle\}$$

It is, of course, not always the case that a constraint can be replaced by a collection of binary constraints on the same variables. In many cases, the binary projections of the constraint relation allow extra solutions, as the following example demonstrates.

Example 5.2.3. Recall the relation δ on domain $D = \{0, 1\}$ defined in Example 2.1.5. The binary projections of δ are as follows:

$$\pi_{1,2}(\delta) = \{\langle 0, 0\rangle, \langle 0, 1\rangle, \langle 1, 0\rangle, \langle 1, 1\rangle\}$$

$$\pi_{1,3}(\delta) = \{\langle 0, 0\rangle, \langle 0, 1\rangle, \langle 1, 0\rangle, \langle 1, 1\rangle\}$$

$$\pi_{2,3}(\delta) = \{\langle 0, 0\rangle, \langle 0, 1\rangle, \langle 1, 0\rangle, \langle 1, 1\rangle\}.$$

The join of these binary projections contains the tuples $\langle 0, 0, 0\rangle$ and $\langle 1, 1, 1\rangle$, which are not elements of δ. It clearly follows that a constraint with relation δ cannot be replaced by any set of binary constraints on the same variables without either losing solutions or introducing extra solutions.

THEOREM 5.2.4. *Let Γ be any set of relations over a finite domain, D.*

If Γ is closed under a majority operation, then CSP(Γ) is solvable in polynomial time.

PROOF. For any problem instance $\mathcal{P}$ in CSP(Γ) we can impose strong $(|D| + 1)$-consistency [Dechter 1992] in polynomial time to obtain a new instance $\mathcal{P}'$ with the same solutions. All of the constraints in $\mathcal{P}'$ are elements of Γ^+, and so they are closed under a majority operation, by Proposition 3.5. Hence, all of the constraints of $\mathcal{P}'$ are decomposable into binary constraints by Proposition 5.2.1. Hence, by Corollary 3.2 of Dechter [1992], $\mathcal{P}'$ is solvable in polynomial time. $\square$

Example 5.2.5. When $D = \{\mathbf{true,false}\}$, there is only one possible majority operation on D, (which is equal to the Δ operation defined in Example 3.4). It is easily shown that all possible binary Boolean relations are closed under Δ. Hence, it follows from Proposition 5.2.1 that the Boolean relations of arbitrary arity, which are closed under this majority operation, are precisely the relations that are definable by a formula in conjunctive normal form in which each conjunct contains at most two literals. Hence, if a set of Boolean relations Γ is closed under a majority operation, then CSP(Γ) is equivalent to the 2-SATISFIABILITY problem (2-SAT) [Papadimitriou 1994], which is well known to be a tractable subproblem of the SATISFIABILITY problem [Schaefer 1978].

Recall the class of tractable constraints identified independently in Cooper et al. [1994] and Kirousis [1993], and referred to as *0/1/all constraints* or *implicational constraints*. (This class of tractable constraints is referred to as *Class I* in Jeavons et al. [1995].) It was shown in Jeavons et al. [1995] that these constraints are in fact precisely the relations closed under the majority operation Δ defined in Example 3.4. This result is rather unexpected, in view of the fact that 0/1/all constraints were originally defined purely in terms of their syntactic structure [Cooper et al. 1994].

However, we remark here that the class of tractable sets of relations defined by closure under some majority operation is a true generalization of the class containing all sets of 0/1/all constraints. In other words, there exist tractable sets of relations that are closed under some majority operation but are not closed under the Δ operation, as the following example demonstrates.

Example 5.2.6. Let μ be the ternary majority operation on $D = \{0, 1, 2\}$ that returns the median value of its three arguments (in the standard ordering of D).

Recall the relation R_3 defined in Example 2.1.6. It is easy to show that R_3 is closed under μ, since applying the μ operation to any three elements of R_3 yields an element of R_3. For example,

$$\mu(\langle 1, 0\rangle, \langle 0, 2\rangle, \langle 0, 1\rangle) = \langle 0, 1\rangle \in R_3.$$

Hence, by Theorem 5.2.4, CSP($\{R_3\}$) is tractable.

However, it was shown in Example 3.4 that R_3 is not closed under Δ, and hence R_3 is not a 0/1/all constraint.

5.3. BINARY OPERATIONS. We first show that closure under an *arbitrary* idempotent binary operation is *not* in general sufficient to ensure tractability.

LEMMA 5.3.1. *There exists a set of relations Γ closed under an idempotent binary operation* (*which is not a projection*) *such that CSP*(Γ) *is NP-complete*.

PROOF. Consider the binary operation $\Box$ on the set $D = \{0, 1, 2, 3\}$, which is defined by the following table:

$\Box$	0	1	2	3
0	0	1	0	1
1	0	1	0	1
2	2	3	2	3
3	2	3	2	3

This operation is idempotent but it is not a projection. (In fact, it is an example of a form of binary operation known as a "rectangular band" [McKenzie et al. 1987].)

Now consider the functions $b_1: D \to \{0, 1\}$ and $b_2: D \to \{0, 1\}$ which return the first and second bit in the binary expression for the numerical value of each element of D.

Using these functions, we define ternary relations R_1 and R_2 over D, as follows:

$$R_1 = \{\langle d, d', d''\rangle \in D^3 | (b_1(d) \neq b_1(d')) \lor (b_1(d') \neq b_1(d'')) \lor (b_1(d'') \neq b_1(d))\}$$

$$R_2 = \{\langle d, d', d''\rangle \in D^3 | (b_2(d) \neq b_2(d')) \lor (b_2(d') \neq b_2(d'')) \lor (b_2(d'') \neq b_2(d))\}.$$

Finally, we define $R = R_1 \cap R_2$.

It is easily shown that R is closed under $\Box$, since applying the operation $\Box$ to any 2 elements of R yields an element of R.

However, it can also be shown that the NOT-ALL-EQUAL SATISFIABILITY problem [Schaefer 1978], which is known to be NP-complete, is reducible in polynomial time to CSP($\{R\}$). Hence, CSP($\{R\}$) is NP-complete, and the result follows. $\Box$

We now describe some additional conditions that may be imposed on binary operations. It will be shown below that closure under any binary operation satisfying these additional conditions is a sufficient condition for tractability.

Definition 5.3.2. Let $\sqcap: D^2 \to D$ be an idempotent binary operation on the set D such that, for all $d_1, d_2, d_3 \in D$,

—$\sqcap(\sqcap(d_1, d_2), d_3) = \sqcap(d_1, \sqcap(d_2, d_3))$; and (*Associativity*)
—$\sqcap(d_1, d_2) = \sqcap(d_2, d_1)$. (*Commutativity*)

Then $\sqcap$ is said to be an *ACI operation.*

We will make use of the following result about ACI operations, which is well known from elementary algebra [Cohn 1965; McKenzie et al. 1987].

LEMMA 5.3.3. *Let $\sqcap$ be an ACI operation on the set D. The binary relation R on D defined by*

$$R(d_1, d_2) \Leftrightarrow \sqcap(d_1, d_2) = d_2$$

is a partial order on D in which any two elements d_1, d_2 have a least upper bound given by $\sqcap(d_1, d_2)$.

It follows from Lemma 5.3.3 that any (finite) nonempty set $D' \subseteq D$ which is $\sqcap$-closed contains a least upper bound with respect to the partial order R. This upper bound will be denoted $\sqcap(D')$.

Using Lemma 5.3.3, we now show that relations that are closed under some arbitrary ACI operation form a tractable class.

THEOREM 5.3.4. *For any set of relations Γ over a finite domain D, if Γ is closed under some ACI operation, then CSP(Γ) is solvable in polynomial time.*

PROOF. Let Γ be a set of relations closed under the ACI operation $\sqcap$, and let $\mathcal{P}$ be any problem instance in CSP(Γ). First enforce pairwise consistency to obtain a new instance $\mathcal{P}'$ with the same set of solutions which is pairwise consistent. Such a $\mathcal{P}'$ can be obtained by forming the join of every pair of constraints in $\mathcal{P}$, replacing these constraints with the (possibly smaller) constraints obtained by projecting down to the original scopes, and then repeating this process until there are no further changes in the constraints. The time complexity of this procedure is polynomial in the size of $\mathcal{P}$, and the resulting $\mathcal{P}'$ is a member of CSP(Γ^+). Hence, all the constraint relations in $\mathcal{P}'$ are closed under $\sqcap$, by Proposition 3.5.

Now let $D(v)$ denote the set of values allowed for variable v by the constraints of $\mathcal{P}'$. Since $D(v)$ equals the projection of some $\sqcap$-closed constraint onto v, it must be $\sqcap$-closed, by Proposition 3.5. There are two cases to consider:

(1) If any of the sets $D(v)$ is empty, then $\mathcal{P}'$ has no solutions, so the decision problem is trivial.
(2) On the other hand, if all of these sets are nonempty, then we claim that assigning the value $\sqcap(D(v))$ to each variable v gives a solution to $\mathcal{P}'$, so the decision problem is again trivial. To establish this claim, consider any constraint $C = (S, R)$ in $\mathcal{P}'$, with relation R of arity n, and scope S. For each $i \in \{1, 2, \ldots, n\}$, there must be some tuple $t_i \in R$ such that $t_i[i] = \sqcap(D(S[i]))$, by the definition of $D(S[i])$. Now consider the tuple $t = \sqcap(t_1, \sqcap(t_2, \ldots, \sqcap(t_{r-1}, t_r))\cdots)$. We know that $t \in R$, since R is closed under $\sqcap$. Furthermore, for each i, $\sqcap(D(S[i]))$ is an upper bound of $D(S[i])$, so

$$\sqcap(d, \sqcap(D(S[i]))) = \sqcap(D(S[i]))$$

for all $d \in D(S[i])$. Hence, $t[i] = \sqcap(D(S[i]))$, so the constraint C allows the assignment of $\sqcap(D(v))$ to each variable v in S. Since C was arbitrary, we have shown that this assignment is a solution to $\mathcal{P}'$, and hence a solution to $\mathcal{P}$. □

Example 5.3.5. When $D = \{\textbf{true,false}\}$, there are only two idempotent binary operations on D (which are not projections), corresponding to the logical AND operation and the logical OR operation. These two operations are both ACI operations, and they correspond to the two possible orderings of D.

It is well known [Dechter and Pearl 1992; Jeavons and Cooper 1996; Papadimitriou 1994] that a Boolean relation is closed under AND if and only if it can be defined by a Horn sentence (i.e., a conjunction of clauses each of which contains at most one unnegated literal). Hence, if a set of Boolean relations Γ is closed under AND, then CSP(Γ) is equivalent to the Horn clause satisfiability problem, HORNSAT [Papadimitriou 1994], which is a tractable subproblem of the SATISFIABILITY problem [Schaefer 1978].

Similarly, a Boolean relation is closed under OR if and only if it can be defined by a conjunction of clauses each of which contains at most one negated literal, and this class or relations also gives rise to a tractable subproblem of the SATISFIABILITY problem [Schaefer 1978].

Example 5.3.6. Let D be a finite subset of the natural numbers. The operation MAX: $D^2 \rightarrow D$ which returns the larger of any pair of numbers is an ACI operation. The following types of arithmetic constraints (amongst many others) are closed under this operation:

—$aX \neq b$

—$aX = bY + c$

—$aX \leq bY + c$

—$aX \geq bY + c$

—$a_1X + a_2Y + \cdots + a_rZ \geq c$

—$aXY \geq c$

—$(a_1X \geq b_1) \vee (a_2Y \geq b_2) \vee (a_3Z \leq b_3)$

where upper-case letters represent variables and lower-case letters represent positive constants. Hence, by Theorem 5.3.4, it is possible to determine efficiently whether any collection of constraints of these types has a solution. These constraints include (and extend) the 'basic' arithmetic constraints allowed by the well-known constraint programming language, CHIP [Van Hentenryck et al. 1992].

The class of tractable constraints first identified in Jeavons and Cooper [1996], and referred to as *max-closed constraints*, are in fact relations closed under an $\sqcap$ operation with the additional property that the partial order R, defined in Lemma 5.3.3, is a total ordering of D. Hence, a set of constraints is max-closed if and only if the constraint relations are closed under some specialized ACI operation of this kind (see, e.g., Example 5.3.6). This class of tractable constraints is referred to as *Class II* in Jeavons et al. [1995].

However, we remark here that the class of tractable sets of relations defined by closure under some ACI operation is a true generalization of the class containing all sets of max-closed constraints. In other words, there exist tractable relations which are closed under some ACI operation but are not closed under the maximum operation associated with any (total) ordering of the domain. (An example of such a relation is the relation R_1 defined in Example 2.1.6, see Example 6.4 below.)

5.4. AFFINE OPERATIONS. We will now show that closure under an affine operation is a sufficient condition for tractability.

This result was established in Jeavons et al. [1995] using elementary methods, for the special case when the domain D contains a prime number, p, of elements. It was shown in Jeavons et al. [1995] that, in this special case, constraints which are closed under an affine operation correspond precisely to constraints which may be expressed as conjunctions of *linear equations* modulo p. (This class of tractable constraints is referred to as *Class III* in Jeavons et al. [1995].)

We now generalize this result to arbitrary finite domain sizes by making use of a result stated by Feder and Vardi [1993].

THEOREM 5.4.1 [FEDER AND VARDI 1993]. *For any finite group G, and any set Γ of cosets of subgroups of direct products of G, $CSP(\Gamma)$ is solvable in polynomial time.*

COROLLARY 5.4.2. *For any set of relations* Γ, *if* Γ *is closed under an affine operation*, *then CSP*(Γ) *is solvable in polynomial time*.

PROOF. By Definitions 3.1 and 3.2, any relation R that is closed under an affine operation is a subset of a direct product of some Abelian group, with the property that for all $t_1, t_2, t_3 \in R$, $t_1 - t_2 + t_3 \in R$. However, this is equivalent to saying that R is a coset of a subgroup of this direct product group [McKenzie et al. 1987], so we may apply Theorem 5.4.1 to Γ to obtain the result. □

Example 5.4.3. Let ∇ be the affine operation on $D = \{0, 1, 2\}$ which is defined by $\nabla(d_1, d_2, d_3) = d_1 - d_2 + d_3$, where addition and subtraction are both modulo 3.

The relation R_2 defined in Example 2.1.6 is closed under ∇, since applying the ∇ operation to any three elements of R_2 yields an element of R_2. For example,

$$\nabla(\langle 0, 1, 2\rangle, \langle 1, 2, 0\rangle, \langle 2, 0, 1\rangle) = \langle 1, 2, 0\rangle \in R_2.$$

Since $|D|$ is prime, the results of Jeavons et al. [1995] indicate that R_2 must be the set of solutions to some system of linear equations over the integers modulo 3. In fact, we have

$$R_2 = \{\langle x_1, x_2, x_3\rangle | (x_1 - x_2 \equiv 2(\text{mod } 3)) \ \wedge \ (x_2 - x_3 \equiv 2(\text{mod } 3))\}.$$

Example 5.4.4. Let G be the Abelian group $\langle D, +, -\rangle$, where $D = \{0, 1, 2, 3\}$ and the $+$ operation is defined by the following table:

+	0	1	2	3
0	0	1	2	3
1	1	0	3	2
2	2	3	0	1
3	3	2	1	0

Now let ∇ be the affine operation on $D = \{0, 1, 2, 3\}$ which is defined by $\nabla(d_1, d_2, d_3) = d_1 - d_2 + d_3$, where addition and subtraction are as defined in G.

Any relation R over D which is a coset of a subgroup of a direct product of G will be closed under ∇, and hence CSP($\{R\}$) will be tractable by Corollary 5.4.2. One example of such a relation is the following:

$$R = \{\langle 0, 1, 2\rangle, \langle 1, 1, 2\rangle, \langle 2, 2, 0\rangle, \langle 3, 2, 0\rangle\}.$$

It is easily seen that in this case R is *not* the set of solutions to any system of linear equations over a field.

Example 5.4.5. When $D = \{\textbf{true,false}\}$, there is only one possible Abelian group structure over D and hence only one possible affine operation on D.

If a set of Boolean relations Γ is closed under this affine operation, then CSP(Γ) is equivalent to the problem of solving a set of simultaneous linear equations over the integers modulo 2. This corresponds to the final tractable subproblem of the GENERALIZED SATISFIABILITY problem identified by Schaefer [1978].

5.5. SEMIPROJECTIONS. We now show that closure under a semiprojection operation is *not* in general a sufficient condition for tractability. In fact, we shall establish a much stronger result, which shows that even being closed under *all* semiprojections is not sufficient to ensure tractability.

LEMMA 5.5.1. *For any finite set D, with $|D| \geq 3$, there exists a set of relations Γ over D, such that Γ is closed under all semiprojections on D, and CSP(Γ) is NP-complete.*

PROOF. Let D be a finite set with $|D| \geq 3$ and let d_1, d_2 be elements of D. Consider the relation $R = \{\langle d_1, d_1, d_2\rangle, \langle d_1, d_2, d_1\rangle, \langle d_2, d_1, d_1\rangle, \langle d_2, d_2, d_1\rangle, \langle d_2, d_1, d_2\rangle, \langle d_1, d_2, d_2\rangle\}$. This relation is closed under all semiprojections on D, since any 3 elements of R contain at most two distinct values in each coordinate position, so the semiprojection reduces to a projection.

However, if we identify d_1 with the Boolean value **true** and d_2 with the Boolean value **false**, then it is easy to see that CSP($\{R\}$) is isomorphic to the NOT-ALL-EQUAL SATISFIABILITY problem [Garey and Johnson 1979], which is NP-complete [Schaefer 1978] (see Example 2.1.5). □

It is currently unknown whether there are tractable sets of relations closed under some combination of semiprojections, unary operations and binary operations that are not included in any of the tractable classes listed above.

However, when $|D| = 2$ the situation is very simple, as the next example shows.

Example 5.5.2. When $D = \{\mathbf{true}, \mathbf{false}\}$, there are no semiprojections on D, so there are no subproblems of the SATISFIABILITY problem that are characterized by a closure operation of this form.

6. *Calculating Closure Operations*

For any set of relations Γ, over a set D, the operations under which Γ is closed are simply mappings from D^k to D, for some k, which satisfy certain constraints, as described in Definition 3.3. In this section, we show that it is possible to identify these operations by solving a single constraint satisfaction problem in CSP(Γ). In fact, we shall show that these closure operations are precisely the solutions to a constraint satisfaction problem of the following form:

Definition 6.1. Let Γ be a set of relations over a finite domain D.

For any natural number $m > 0$, the *indicator problem* for Γ of order m is defined to be the constraint satisfaction problem $\mathcal{IP}(\Gamma, m)$ with

—Set of variables D^m;

—Domain of values D;

—Set of constraints $\{C_1, C_2, \ldots, C_q\}$, such that for each $R \in \Gamma$, and for each sequence $t_1, t_2, \ldots, t_m$ of tuples from R, there is a constraint $C_i = (S_i, R)$ with $S_i = (v_1, v_2, \ldots, v_n)$ where n is the arity of R and $v_j = \langle t_1[j], t_2[j], \ldots, t_m[j]\rangle$.

Example 6.2. Consider the relation R_1 over $D = \{0, 1, 2\}$, defined in Example 2.1.6.

The indicator problem for $\{R_1\}$ of order 1, $\mathcal{IP}(\{R_1\}, 1)$, has three variables and four constraints. The set of variables is

$$\{\langle 0\rangle, \langle 1\rangle, \langle 2\rangle\},$$

and the set of constraints is

$$\begin{aligned}\{\quad & ((\langle 0\rangle, \langle 0\rangle), R_1),\\ & ((\langle 0\rangle, \langle 1\rangle), R_1),\\ & ((\langle 1\rangle, \langle 2\rangle), R_1),\\ & ((\langle 2\rangle, \langle 1\rangle), R_1) \quad \}.\end{aligned}$$

The indicator problem for $\{R_1\}$ of order 2, $\mathcal{IP}(\{R_1\}, 2)$, has nine variables and 16 constraints. The set of variables is

$$\{\langle 0, 0\rangle, \langle 0, 1\rangle, \langle 0, 2\rangle, \langle 1, 0\rangle, \langle 1, 1\rangle, \langle 1, 2\rangle, \langle 2, 0\rangle, \langle 2, 1\rangle, \langle 2, 2\rangle\},$$

and the set of constraints is

$$\begin{aligned}\{\quad & ((\langle 0, 0\rangle, \langle 0, 0\rangle), R_1), && ((\langle 0, 0\rangle, \langle 0, 1\rangle), R_1),\\ & ((\langle 0, 0\rangle, \langle 1, 0\rangle), R_1), && ((\langle 0, 0\rangle, \langle 1, 1\rangle), R_1),\\ & ((\langle 0, 1\rangle, \langle 0, 2\rangle), R_1), && ((\langle 0, 1\rangle, \langle 1, 2\rangle), R_1),\\ & ((\langle 0, 2\rangle, \langle 0, 1\rangle), R_1), && ((\langle 0, 2\rangle, \langle 1, 1\rangle), R_1),\\ & ((\langle 1, 0\rangle, \langle 2, 0\rangle), R_1), && ((\langle 1, 0\rangle, \langle 2, 1\rangle), R_1),\\ & ((\langle 1, 1\rangle, \langle 2, 2\rangle), R_1), && ((\langle 1, 2\rangle, \langle 2, 1\rangle), R_1),\\ & ((\langle 2, 0\rangle, \langle 1, 0\rangle), R_1), && ((\langle 2, 0\rangle, \langle 1, 1\rangle), R_1),\\ & ((\langle 2, 1\rangle, \langle 1, 2\rangle), R_1), && ((\langle 2, 2\rangle, \langle 1, 1\rangle), R_1) \quad \}.\end{aligned}$$

Further illustrative examples of indicator problems are given in Jeavons et al. [1996].

Solutions to the indicator problem for Γ of order m are functions from D^m to D, or in other words, m-ary operations on D. We now show that they are precisely the m-ary operations under which Γ is closed.

THEOREM 6.3. *For any set of relations Γ over domain D, the set of solutions to $\mathcal{IP}(\Gamma, m)$ is equal to the set of m-ary operations under which Γ is closed.*

PROOF. By Definition 3.3, we know that Γ is closed under the m-ary operation $\otimes$ if and only if $\otimes$ satisfies the condition $\otimes(t_1, t_2, \ldots, t_m) \in R$ for each possible choice of $R \in \Gamma$ and $t_1, t_2, \ldots, t_m \in R$ (not necessarily all distinct). But, by Definition 6.1, this is equivalent to saying that $\otimes$ satisfies all the constraints in $\mathcal{IP}(\Gamma, m)$, so the result follows. $\square$

Example 6.4. Consider the relation R_1 over $D = \{0, 1, 2\}$, defined in Example 2.1.6.

The indicator problem for $\{R_1\}$ of order 1, defined in Example 6.2, has two solutions, which may be expressed in tabular form as follows:

	Variables		
	$\langle 0\rangle$	$\langle 1\rangle$	$\langle 2\rangle$
Solution 1	0	0	0
Solution 2	0	1	2

One of these solutions is a constant operation, so CSP($\{R_1\}$) is tractable, by Proposition 5.1.1. In fact, any problem in CSP($\{R_1\}$) has the solution that assigns the value 0 to each variable, so the complexity of CSP($\{R_1\}$) is trivial.

The indicator problem for $\{R_1\}$ of order 2, defined in Example 6.2, has 4 solutions, which may be expressed in tabular form as follows:

	Variables								
	$\langle 0,0\rangle$	$\langle 0,1\rangle$	$\langle 0,2\rangle$	$\langle 1,0\rangle$	$\langle 1,1\rangle$	$\langle 1,2\rangle$	$\langle 2,0\rangle$	$\langle 2,1\rangle$	$\langle 2,2\rangle$
Solution 1	0	0	0	0	0	0	0	0	0
Solution 2	0	1	2	0	1	2	0	1	2
Solution 3	0	0	0	1	1	1	2	2	2
Solution 4	0	0	0	0	1	0	0	0	2

The first of these solutions is a constant operation, and the second and third are essentially unary operations. However, the fourth solution shown in the table is more interesting. It is easily checked that this operation is an associative, commutative, idempotent (ACI) binary operation, so we have a second proof that CSP($\{R_1\}$) is tractable, by Theorem 5.3.4. Furthermore, this result shows that R_1 can be combined with any other relations (of any arity) that are also closed under this ACI operation to obtain larger tractable problem classes.

COROLLARY 6.5. *For any set of relations* Γ *over a domain* D, *with* $|D| \geq 3$, *if all solutions to* $\mathcal{IP}(\Gamma, |D|)$ *are essentially unary, then CSP*(Γ) *is NP-complete.*

PROOF. Follows from Theorem 3.8, Theorem 4.1, and Theorem 6.3. □

Example 6.6. Recall the relations R_1, R_2, and R_3 defined in Example 2.1.6. It has been shown in Examples 6.4, 5.4.3, and 5.1.2 that a set containing any *one* of these relations on its own is tractable.

For any set Γ containing more than one of these relations, it can be shown, using Corollary 6.5, that CSP(Γ) is NP-complete.

In the special case when $|D| = 2$, we obtain an even stronger result.

COROLLARY 6.7. *For any set of relations* Γ *over a domain* D, *with* $|D| = 2$, *if all solutions to* $\mathcal{IP}(\Gamma, 3)$ *are essentially unary, then CSP*(Γ) *is NP-complete*; *otherwise*, *it is polynomial.*

PROOF. It has been shown in Examples 5.1.3, 5.2.5, 5.3.5, 5.4.5, and 5.5.2 that, when $|D| = 2$, all possible closure operations of the restricted types specified in Corollary 4.2 are sufficient to ensure tractability. □

This result demonstrates that solving the indicator problem of order 3 provides a simple and complete test for tractability of any set of relations over a domain with 2 elements. This answers a question posed by Schaefer [1978] concerning the existence of an efficient test for tractability in the GENERALIZED SATISFIABILITY problem. Note that carrying out the test requires finding the solutions to a constraint satisfaction problem with just eight Boolean variables.

7. *Conclusion*

In this paper, we have shown how the algebraic properties of relations can be used to distinguish between sets of relations that give rise to tractable constraint satisfaction problems and those which give rise to NP-complete problems. Furthermore, we have proposed a method for determining the operations under which a set of relations is closed by solving a particular form of constraint satisfaction problem, which we have called an indicator problem.

For problems where the domain contains just two elements, these results provide a necessary and sufficient condition for tractability (assuming that P is not equal to NP), and an efficient test to distinguish the tractable sets of relations.

For problems with larger domains, we have described algebraic closure properties that are a necessary condition for tractability. We have also shown that, in many cases, these closure properties are sufficient to ensure tractability.

In particular, we have shown that closure under any constant operation, any majority operation, any ACI operation, or any affine operation, is a sufficient condition for tractability. It can be shown using the results of Jeavons [1995] that for any operation of one of these types, the set, Γ, containing all relations which are closed under that operation is a *maximal* set of tractable relations. In other words, the addition of *any* other relation which is not closed under the same operation changes CSP(Γ) from a tractable problem into an NP-complete problem. Hence, the tractable classes defined in this way are as large as possible.

We are now investigating the application of these results to particular problem types, such as temporal problems involving subsets of the interval algebra. We are also attempting to determine how the presence of particular algebraic closure properties in the constraints can be used to derive appropriate efficient algorithms for tractable problems.

ACKNOWLEDGMENTS. We are grateful to Martin Cooper for many helpful discussions, and for suggesting the proof of Theorem 5.2.4.

REFERENCES

CODD, E. F. 1970. A relational model of data for large shared data banks. *Commun. ACM 13*, 6 (June), 377–387.

COHEN, D., GYSSENS, M., AND JEAVONS, P. 1996. Derivation of constraints and database relations. In *Proceedings of the 2nd International Conference on Constraint Programming—CP '96* (Boston, Mass., Aug.). Lecture Notes in Computer Science, 1118. Springer-Verlag, New York, pp. 134–148.

COHN, P. M. 1965. *Universal Algebra*. Harper & Row, New York.

COOPER, M. C., COHEN, D. A. AND JEAVONS, P. G. 1994. Characterizing tractable constraints. *Artif. Int. 65*, 347–361.

DECHTER, R. 1992. From local to global consistency. *Artif. Int. 55*, 87–107.

DECHTER, R., AND PEARL, J. 1988. Network-based heuristics for constraint-satisfaction problems. *Artif. Int. 34*, 1–38.

DECHTER, R., AND PEARL, J. 1992. Structure identification in relational data. *Artif. Int. 58*, 237–270.

FEDER, T., AND VARDI, M. 1993. Monotone monadic SNP and constraint satisfaction. In *Proceedings of the 25th ACM Symposium on the Theory of Computing* (*STOC*) (San Diego, Calif., May 16–18). ACM, New York, pp. 612–622.

FREUDER, E. C. 1985. A sufficient condition for backtrack-bounded search. *J. ACM 32*, 4 (Oct.), 755–761.

FREUDER, E. C. 1991. Eliminating interchangeable values in constraint satisfaction problems. In *Proceedings of AAAI-91*. pp. 227–233.

GAREY, M. R., AND JOHNSON, D. S. 1979. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, San Francisco, Calif.

GYSSENS, M., JEAVONS, P., AND COHEN, D. 1994. Decomposing constraint satisfaction problems using database techniques. *Artif. Int. 66*, 57–89.

JEAVONS, P. G. 1995. An algebraic characterization of tractable constraints. Tech. Rep. CSD-TR-95-05. Royal Holloway, Univ. London, London, United Kingdom.

JEAVONS, P., COHEN, D., AND GYSSENS, M. 1995. A unifying framework for tractable constraints. In *Proceedings of the 1st International Conference on Principles and Practice of Constraint Programming—CP '95* (Cassis, France, Sept.). Lecture Notes in Computer Science, vol. 976. Springer-Verlag, New York, pp. 276–291.

JEAVONS, P., COHEN, D., AND GYSSENS, M. 1996. A test for tractability. In *Proceedings of the 2nd International Conference on Constraint Programming—CP '96* (Boston, Mass., Aug.). Lecture Notes in Computer Science, vol. 1118. Springer-Verlag, New York, pp. 267–281.

JEAVONS, P. G., AND COOPER, M. C. 1996. Tractable constraints on ordered domains. *Artif. Int. 79*, 327–339.

KIROUSIS, L. 1993. Fast parallel constraint satisfaction. *Artif. Int. 64*, 147–160.

LADKIN, P. B., AND MADDUX, R. D. 1994. On binary constraint problems. *J. ACM 41*, 3 (May), 435–469.

MACHIDA, H., AND ROSENBERG, I. G. 1984. Classifying essentially minimal clones. In *Proceedings of the 14th International Symposium on Multiple-Valued Logic*. (Winnipeg, 1984), pp. 4–7.

MACKWORTH, A. K. 1977. Consistency in networks of relations. *Artif. Int. 8*, 99–118.

MCKENZIE, R. N., MCNULTY, G. F., AND TAYLOR, W. F. 1987. *Algebras, Lattices and Varieties. Volume I*. Wadsworth and Brooks, Calif.

MONTANARI, U. 1974. Networks of constraints: Fundamental properties and applications to picture processing. *Inf. Sci. 7*, 95–132.

MONTANARI, U., AND ROSSI, F. 1991. Constraint relaxation may be perfect. *Artif. Int. 48*, 143–170.

PAPADIMITRIOU, C. H. 1994. *Computational Complexity*, Addison-Wesley, Reading, Pa.

ROSENBERG, I. G. 1986. Minimal clones I: The five types. In *Lectures in Universal Algebra* (*Proc. Conf. Szeged 1983*), *Colloq. Math. Soc. Janos Bolyai*, vol. 43. North-Holland, Amsterdam, The Netherlands, pp. 405–427.

SCHAEFER, T. J. 1978. The complexity of satisfiability problems. In *Proceedings of the 10th ACM Symposium on Theory of Computing* (San Diego, Calif., May 1–3). ACM, New York, pp. 216–226.

SZENDREI, A. 1986. *Clones in Universal Algebra*. Seminaires de Mathematiques Superieures 99, University of Montreal, Montreal, Que., Canada.

VAN BEEK, P. 1992. On the minimality and decomposability of row-convex constraint networks. In *Proceedings of the 10th National Conference on Artificial Intelligence*, *AAAI-92*. MIT Press, Cambridge, Mass., pp. 447–452.

VAN HENTENRYCK, P., DEVILLE, Y., AND TENG, C-M. 1992. A generic arc-consistency algorithm and its specializations. *Artif. Int. 57*, 291–321.

RECEIVED OCTOBER 1996; REVISED MAY 1997; ACCEPTED APRIL 1997