

Robert COWEN

A COMPACTNESS THEOREM FOR INFINITE CONSTRAINT SATISFACTION

A b s t r a c t. A useful compactness theorem for constraint satisfaction problems is proved equivalent to BPI, the Boolean Prime Ideal Theorem. The relation of various restricted versions of the Theorem to each other and to BPI is also explored.

1. Introduction

Constraint satisfaction problems are very common in computer science; in fact it has been claimed in [2] that most of the problems in Garey and Johnson [5] can be naturally expressed as constraint satisfaction. Since many of these problems have infinite analogues it seems reasonable to consider infinite constraint satisfaction. Then by proving a compactness result for the infinite constraint satisfaction, a useful general theorem is obtained which gives immediately various well known compactness results for propositional logic, graph coloring, etc.

Received December 4, 1998

2. Finite constraint satisfaction

Diverse problems such as satisfiability of a finite set of propositional formulas, solving a finite set of equations over a finite field, choosing a set of distinct representatives for a collection of finite sets (marriage problem), vertex coloring a finite graph with k colors, etc., can all be expressed in the following format. Given a set D and a set of variables V , each one associated with a finite set of ‘allowable’ values in D so that all assignments made to the variables must select only from the allowable values for each variable. Can an assignment to the variables be made, such that the conjunction,

$$\bigwedge_{1 \leq i \leq m} R_i(x_{i_1}, \dots, x_{i_{k_i}})$$

is satisfied where the x_{i_j} are occurrences of the variables and the R_i are k_i -ary relations on D ? This is the general constraint satisfaction problem (see [2],[8]). We shall refer to conjunctions of the above form as *constraint formulas* or *cfs*.

For example the existence of a k -coloring of a graph G is equivalent to satisfying: $\bigwedge_{\{v,w\} \in E} x_v \neq x_w$, where E is the set of edges of G , and where all variables range over $\{1, \dots, k\}$; however, if each vertex variable is allowed to range over its own list of “colors”, satisfaction is equivalent to a list coloring of the vertices (see West[9] for more on list colorings).

The general constraint satisfaction problem is NP-Complete and remains NP-Complete even if (1) each relation is binary and each variable has only three possible values, or, (2) each relation is ternary and each variable has only two possible values (see [2]); however, it is polynomial if (3) each relation is binary and each variable has only two possible values. In case (3) it is easily shown to be equivalent to 2-SAT, satisfiability of propositional cnfs with only 2 literals per clause which is known to be solvable in polynomial time (see [5]).

3. Infinite constraint satisfaction

All of the problems mentioned above have infinite analogues that can be considered to be special cases of the following problem.

Let I be an index set of arbitrary cardinality. Is there an assignment of allowable values to the variables which satisfies,

$$\bigwedge_{i \in I} R_i(x_{i_1}, \dots, x_{i_{k_i}})$$

where the variables range over *finite* subsets of D and the R_i are k_i -ary relations on D ? To say that an assignment satisfies the above infinite conjunction, merely means that it satisfies all $R_i(x_{i_1}, \dots, x_{i_{k_i}})$, $i \in I$.

The following compactness theorem for constraint satisfaction problems will be referred to as the *Constraint Compactness Theorem* or *CCT*.

Theorem 3.1. (CCT) *Let $\{R_i(x_{i_1}, \dots, x_{i_{k_i}})\}_{i \in I}$ be given, where the R_i represent k_i -ary relations on D and the variables range over finite subsets of D . If $\bigwedge_{i \in W} R_i(x_{i_1}, \dots, x_{i_{k_i}})$ is satisfiable for each finite $W \subset I$, then $\bigwedge_{i \in I} R_i(x_{i_1}, \dots, x_{i_{k_i}})$ is satisfiable.*

It is not hard to show that CCT can be given an equivalent formulation, as follows.

Theorem 3.2. (CCT) *Let $\{\varphi_i\}_{i \in I}$ be a collection of constraint formulas on D and suppose that every finite subset is satisfiable. Then $\{\varphi_i\}_{i \in I}$ is satisfiable.*

The Constraint Compactness Theorem will be proved to be equivalent in Zermelo–Fraenkel Set Theory to the Prime Ideal Theorem for Boolean algebras (BPI). BPI is weaker than the Axiom of Choice but is a most useful Theorem in its own right with many equivalent formulations (see

Howard and Rubin [6]). In fact even restricting CCT by requiring (1) each relation is binary and each variable has only three possible values, or, (2) each relation is ternary and each variable has only two possible values, still gives theorems equipotent to BPI. However, maintaining the analogy with finite constraint satisfaction and NPC, if (3) each relation is binary and each variable has only two possible values, a weaker theorem results.

The claims made above will now be proven. In what follows we regard functions as special sets of ordered pairs and if g is a function and W is a subset of its domain, $g|W$ will denote the restriction of g to W . The following theorem was proved in [3] to be equivalent to BPI in ZF.

Theorem 3.3. *Let H be a set of partial functions on I such that $\{h(i)|h \in H\}$ is a finite set for each $i \in I$. Suppose for each finite $W \subset I$ there is a non empty set $H_W \subset H$ whose domains include W and such that $W_1 \subset W_2$ implies $H_{W_2} \subset H_{W_1}$. Then there exists a function g in H , with domain I , such that for any finite $W \subset I$ there exists $h \in H_W$ with $g|W \subset h$.*

Theorem 3.4. *BPI implies CCT in ZF.*

Proof. We prove CCT from Theorem 3. Assume the hypothesis of CCT. For each $i \in I$, let A_i be the set of assignments to the variables, $\{x_{i_1}, \dots, x_{i_{k_i}}\}$, which satisfy $R_i(x_{i_1}, \dots, x_{i_{k_i}})$. A_i is necessarily finite since the variables range over finite sets. Let H be the set of partial functions on I with the property that $h(i) \in A_i$, for $i \in \text{domain}(h)$. For each finite $W \subset I$, let H_W be those functions h in H , whose domain includes W , with the property that $\cup\{h(i)|i \in W\}$ is a function (which must then satisfy $\bigwedge_{i \in W} R_i(x_{i_1}, \dots, x_{i_{k_i}})$). Then the hypothesis of CCT implies H_W is non empty, since $\bigwedge_{i \in W} R_i(x_{i_1}, \dots, x_{i_{k_i}})$ is satisfiable. Surely $W_1 \subset W_2$ implies $H_{W_2} \subset H_{W_1}$. Thus, BPI (in the form of Theorem 3) gives a function g , with domain I , such that for any finite $W \subset I$ there exists $h \in H_W$ with $g|W \subset h$. We claim that $\cup g$ satisfies $\bigwedge_{i \in I} R_i(x_{i_1}, \dots, x_{i_{k_i}})$. It only remains to show that if $i, j \in I$ then $g(i), g(j)$ agree on their common variables; however, if $W = \{i, j\}$, there exist $h \in H_W$ with $g|W \subset h$. Thus

$g(i) \cup g(j) \subset h(i) \cup h(j)$, a function, which implies $g(i), g(j)$ agree on their common variables. ■

Theorem 3.5. *CCT implies BPI in ZF, even if (a) all relations are binary and all variables have only three possible values or (b) all relations are ternary and all variables have only two possible values.*

Proofs.

(a) It has been proved by Läuchli ([7]) that the compactness for 3-colorable graphs implies BPI. Since, as mentioned above, graph coloring is equivalent to a conjunction of terms of the form $x \neq y$, the theorem easily follows.

(b) 3-SAT is equivalent to BPI, where 3-SAT is the compactness theorem for propositional logic, restricted to formulas in conjunctive normal form, each conjunct of which is a disjunction of at most three literals (a literal is a statement letter or its negation). See [4]. ■

Theorem 3.6. *Restricting CCT by requiring all relations must be binary and all variables have only two possible values results in a theorem which is weaker than BPI in ZF.*

Proof. P. Wojtylak [10] has proved that 2-SAT, the compactness result for propositional clauses with at most two literals each, is weaker than BPI in ZF. He also shows that 2-SAT implies the Axiom of Choice for pairs. We claim that 2-SAT implies CCT restricted by allowing only binary relations and 2-valued variables.

Suppose now that $\{R_i(x_{i_1}, x_{i_2})\}_{i \in I}$ is given, where each variable is allowed only 2 values, and $\{R_i(x_{i_1}, x_{i_2})\}_{i \in W}$ is satisfiable for every finite $W \subset I$. Because we have the Axiom of Choice for pairs, we can assume that each variable ranges over an ordered pair, (a_0, a_1) . For each variable x which appears, we take a new variable v_x which ranges over $\{0, 1\}$.

For any $R(x, y)$ in the indexed set, where x ranges over (a_0, a_1) and y ranges over (b_0, b_1) , we form a cnf, $X_{R(x,y)}$, as follows:

- (1) If $(a_0, b_0) \notin R$, add $(v_x \vee v_y)$ to $X_{R(x,y)}$.

- (2) If $(a_0, b_1) \notin R$, add $(v_x \vee \sim v_y)$ to $X_{R(x,y)}$
- (3) If $(a_1, b_0) \notin R$, add $(\sim v_x \vee v_y)$ to $X_{R(x,y)}$.
- (4) If $(a_1, b_1) \notin R$, add $(\sim v_x \vee \sim v_y)$ to $X_{R(x,y)}$.

Then, as easily shown, $x = a_i, y = b_j$ satisfies $R(x, y)$ if and only if $v_x = i, v_y = j$ satisfies $X_{R(x,y)}$, $0 \leq i, j \leq 1$. We can then replace questions about the satisfiability of $\{R_i(x_{i_1}, x_{i_2})\}_{i \in I}$ and its subsets by equivalent questions about the satisfiability of $\{X_{R_i(x_{i_1}, x_{i_2})}\}_{i \in I}$ and its subsets. Therefore the satisfiability of $\{R_i(x_{i_1}, x_{i_2})\}_{i \in W}$, for finite $W \subset I$ implies the satisfiability of $\{X_{R_i(x_{i_1}, x_{i_2})}\}_{i \in W}$. Then, by 2-SAT, $\{X_{R_i(x_{i_1}, x_{i_2})}\}_{i \in I}$ is satisfiable; but this implies that $\{R_i(x_{i_1}, x_{i_2})\}_{i \in I}$ is satisfiable as well.

Thus 2-SAT implies CCT, restricted to binary relations and 2-valued variables. Since 2-SAT is weaker than BPI, so is this restricted version of CCT. ■

4. Relative Compactness

In the absence of BPI, one cannot, of course, prove CCT in ZF. However, it can often be shown that a class of constraint satisfaction problems is compact (CCT holds for that class) if some other class is assumed to be compact. This “relative compactness” result is often proved by “embedding” one class of problems in the other and then using the assumed compactness of the second class to prove the compactness of the first class. Most often this is used to embed a class whose compactness is provably equivalent to BPI in some subclass thereby showing the compactness of the subclass is equivalent to BPI as well. We shall formalize this method as follows.

If φ is a cf and $var(\varphi)$ is its set of constrained variables, $sat(\varphi)$ will denote the set of those assignments with domain $var(\varphi)$ which satisfy φ . If φ, ψ are cfs with $var(\varphi) \subseteq var(\psi)$, then $\varphi \prec \psi$ shall denote: $sat(\psi)|var(\varphi) = sat(\varphi)$. Thus, if $\varphi \prec \psi$, every satisfying assignment of φ can be extended to a satisfying assignment of ψ and any satisfying assignment of ψ , if restricted to the variables of φ will satisfy φ ; in particular φ

is satisfiable if and only if ψ is satisfiable. It is easily proved that \prec is a transitive relation.

Let \mathcal{R} be a set of relations on D . We shall say, \mathcal{R} is *compact*, and write, $\text{compact}(\mathcal{R})$, if CCT holds when restricted to relations in \mathcal{R} . The following definition is useful in showing that the compactness of one class implies that of another.

Let \mathcal{R} and \mathcal{S} be sets of relations on D . A *conjunctive embedding* from \mathcal{R} to \mathcal{S} is a mapping φ which associates with each $R(x_1, \dots, x_k)$, $R \in \mathcal{R}$, a cf, $\varphi(R(x_1, \dots, x_k)) = \bigwedge_{1 \leq i \leq m} S_i(x_{i_1}, \dots, x_{i_{k_i}})$, $S_i \in \mathcal{S}$, such that, $R(x_1, \dots, x_k) \prec \varphi(R(x_1, \dots, x_k))$. If there is a conjunctive embedding from \mathcal{R} to \mathcal{S} , we shall write $\mathcal{R} \prec \mathcal{S}$. An embedding, φ , from \mathcal{R} to \mathcal{S} will be called *conflict-free* if any variables new to $R(x_1, \dots, x_k)$ introduced in $\varphi(R(x_1, \dots, x_k))$ do not occur in any other $\varphi(T(y_1, \dots, y_k))$. It can always be arranged that embeddings are conflict-free and we shall assume this to be the case in what follows.

The following Theorem is easily proved and hence we shall omit the proof.

Theorem 4.1. *Suppose \mathcal{R} , \mathcal{S} and \mathcal{T} are sets of relations on D . If $\mathcal{R} \prec \mathcal{S}$ and $\mathcal{S} \prec \mathcal{T}$ then $\mathcal{R} \prec \mathcal{T}$.*

Let \mathcal{R}_{sat} stand for all relations on $\{0, 1\}$ which are the satisfying assignments in the sense of propositional logic of some disjunction of literals. For example, $\sim x \vee y$ yields the relation $\{(0, 0), (0, 1), (1, 1)\}$. Let \mathcal{R}_{k-sat} denote those relations in \mathcal{R}_{sat} which come from disjunctions of k literals.

Theorem 4.2. $\mathcal{R}_{sat} \prec \mathcal{R}_{3-sat}$.

Proof. The construction used in Garey and Johnson[5] in proving 3-SAT is NP-complete constitutes a conjunctive embedding of \mathcal{R}_{sat} to \mathcal{R}_{3-sat} (See Cowen[4]). ■

There is no conjunctive embedding of \mathcal{R}_{3-sat} to \mathcal{R}_{2-sat} (and hence no conjunctive embedding of \mathcal{R}_{sat} to \mathcal{R}_{2-sat}) as we show next.

We will prove the following Theorem.

Theorem 4.3. *There is no 2-cnf X with $a \vee b \vee c \prec X$.*

We first prove a lemma.

Lemma 4.4. *Let X be a 2-cnf and suppose that $a_1 \wedge \dots \wedge a_n \wedge X$ is unsatisfiable. Then $a_i \wedge a_j \wedge X$ is unsatisfiable for some i, j where $1 \leq i, j \leq n$.*

Proof. If a is a literal, \bar{a} will denote its opposite. We can assume, without loss of generality, that X is satisfiable and $a_i \neq \bar{a}_j$, for $1 \leq i, j \leq n$, since otherwise the conclusion clearly follows.

We do an induction on the number of clauses in X , that is, let $P(k)$ be the statement of the Theorem when X has k clauses and n is any integer ≥ 2 .

If $X = (x \vee y)$ and $a_1 \wedge \dots \wedge a_n \wedge X$ is unsatisfiable then both $a_1 \wedge \dots \wedge a_n \wedge x$, $a_1 \wedge \dots \wedge a_n \wedge y$ are unsatisfiable. Since $a_i \neq \bar{a}_j$, both $x = \bar{a}_i$, $y = \bar{a}_j$, for some i, j with $1 \leq i, j \leq n$. Hence $a_i \wedge a_j \wedge X$ is unsatisfiable and $P(1)$ is true.

Assume now $P(k)$ holds whenever $k < m$ and suppose X has exactly m clauses. Suppose $X = Z \wedge (x \vee y)$ and suppose X is satisfiable, but $a_1 \wedge \dots \wedge a_n \wedge X$ is unsatisfiable, where $a_i \neq \bar{a}_j$, for $1 \leq i, j \leq n$. Hence Z is satisfiable but both $a_1 \wedge \dots \wedge a_n \wedge x \wedge Z$, $a_1 \wedge \dots \wedge a_n \wedge y \wedge Z$ are unsatisfiable. If $x \neq \bar{a}_i$ and $y \neq \bar{a}_i$, for all i , $1 \leq i \leq n$, the induction hypothesis applies to both $a_1 \wedge \dots \wedge a_n \wedge x \wedge Z$, $a_1 \wedge \dots \wedge a_n \wedge y \wedge Z$; hence, either $a_i \wedge a_j \wedge Z$ is unsatisfiable or both $x \wedge a_i \wedge Z$, $y \wedge a_j \wedge Z$, are unsatisfiable, for some i, j , $1 \leq i, j \leq n$. In the former case, $a_i \wedge a_j \wedge X$ is unsatisfiable. In the latter case, $(x \wedge a_i \wedge Z) \vee (y \wedge a_j \wedge Z)$ is unsatisfiable; therefore $a_i \wedge a_j \wedge X = a_i \wedge a_j \wedge Z \wedge (x \vee y)$ is unsatisfiable, as well.

We can assume then that $x = \bar{a}_i$ or $y = \bar{a}_j$. If both are true, then clearly $a_i \wedge a_j \wedge X$ is unsatisfiable. Suppose exactly one is true, say, $x = \bar{a}_i$, for some i , $1 \leq i \leq n$, but $y \neq \bar{a}_j$, for all j , $1 \leq j \leq n$. Since $x = \bar{a}_i$, $x \wedge a_i \wedge Z$ is clearly unsatisfiable. Since $y \neq \bar{a}_j$, $1 \leq j \leq n$, applying the

induction hypothesis to $a_1 \wedge \dots \wedge a_n \wedge y \wedge Z$ gives either 1) $a_i \wedge a_j \wedge Z$ is unsatisfiable, for some i, j , $1 \leq i, j \leq n$, or 2) $y \wedge a_j \wedge Z$ is unsatisfiable for some j , $1 \leq j \leq n$; in the first case $a_i \wedge a_j \wedge X$ is unsatisfiable. In the second case, since both $y \wedge a_j \wedge Z$ and $x \wedge a_i \wedge Z$ are unsatisfiable, $a_i \wedge a_j \wedge (x \vee y) \wedge Z$ is also unsatisfiable; thus $a_i \wedge a_j \wedge X$ is unsatisfiable. ■

We can now give the proof of the Theorem.

Proof of Theorem 4.3. Suppose $a \vee b \vee c \prec X$; then $\bar{a} \wedge \bar{b} \wedge \bar{c} \wedge X$ is unsatisfiable, since otherwise there would be an interpretation of X whose restriction to $\{a, b, c\}$ fails to satisfy $a \vee b \vee c$. Therefore the lemma implies that at least one of $\bar{a} \wedge \bar{b} \wedge X$, $\bar{a} \wedge \bar{c} \wedge X$, $\bar{b} \wedge \bar{c} \wedge X$ is unsatisfiable. Suppose without loss of generality, $\bar{a} \wedge \bar{b} \wedge X$ is unsatisfiable. Since $a = 0$, $b = 0$, $c = 1$ satisfies $a \vee b \vee c$ and $a \vee b \vee c \prec X$, there is an extension of $a = 0$, $b = 0$, $c = 1$ which satisfies X ; but this implies that $\bar{a} \wedge \bar{b} \wedge X$ is satisfiable! Therefore $a \vee b \vee c \prec X$ must be false. ■

The fact that \mathcal{R}_{sat} is not conjunctively embeddable in \mathcal{R}_{2-sat} , demonstrates the weakness of \mathcal{R}_{2-sat} and, to our mind, suggests that 2-SAT is weaker than BPI in ZF (the result proved by Wojtylak mentioned above). In fact we make the following conjecture.

Conjecture 4.5. *If \mathcal{R} is a set of relations on $\{0, 1\}$ and $\mathcal{R}_{sat} \prec \mathcal{R}$ is false, then $compact(\mathcal{R})$ is weaker than BPI, in ZF.*

However, if $\mathcal{R}_{sat} \prec \mathcal{R}$ is true, it does follow that $compact(\mathcal{R})$ is equivalent to BPI. This follows from the following ‘‘Relative Compactness Theorem.’’

Theorem 4.6. *Let \mathcal{R} and \mathcal{S} be sets of relations on D suppose that $\mathcal{R} \prec \mathcal{S}$. Then $compact(\mathcal{S})$ implies $compact(\mathcal{R})$.*

Proof. Assume $compact(\mathcal{S})$ and that φ is a conjunctive, conflict-free embedding from \mathcal{R} to \mathcal{S} . Let $\{R_i(x_{i_1}, \dots, x_{i_{k_i}})\}_{i \in I}$ be given where the R_i represent relations on D and the variables range over finite subsets of D .

Suppose $\bigwedge_{i \in W} R_i(x_{i_1}, \dots, x_{i_{k_i}})$ is satisfiable for each finite $W \subset I$; we must show that $\bigwedge_{i \in I} R_i(x_{i_1}, \dots, x_{i_{k_i}})$ is satisfiable.

Let $\Sigma = \{S_j(y_{j_1}, \dots, y_{k_j})\}_{j \in J}$, where the $S_j(y_{j_1}, \dots, y_{k_j})$ occur in the $\varphi(R_i(x_{i_1}, \dots, x_{i_{k_i}}))$, $i \in I$. We claim that Σ is satisfiable. Suppose that $W \subset J$, W finite. Let $F(W)$ be a finite set such that each $S_j(y_{j_1}, \dots, y_{k_j})$, $j \in W$, occurs in some $\varphi(R_i(x_{i_1}, \dots, x_{i_{k_i}}))$, $i \in F(W)$. (Since W is finite, the Axiom of Choice is not needed here.) Since $\bigwedge_{i \in F(W)} R_i(x_{i_1}, \dots, x_{i_{k_i}})$ is satisfiable and φ is conflict-free, $\bigwedge_{i \in F(W)} \varphi(R_i(x_{i_1}, \dots, x_{i_{k_i}}))$ is satisfiable as well. But then, surely, $\bigwedge_{j \in W} S_j(y_{j_1}, \dots, y_{k_j})$ will also be satisfiable. Since \mathcal{S} is compact, $\bigwedge_{j \in J} S_j(y_{j_1}, \dots, y_{k_j})$ is satisfiable. Hence, $\bigwedge_{i \in I} \varphi(R_i(x_{i_1}, \dots, x_{i_{k_i}}))$ is satisfiable, and finally, the restriction of any satisfying assignment to the variables of the $R_i(x_{i_1}, \dots, x_{i_{k_i}})$, $i \in I$, must satisfy $\bigwedge_{i \in I} R_i(x_{i_1}, \dots, x_{i_{k_i}})$, since $R_i(x_{i_1}, \dots, x_{i_{k_i}}) \prec \varphi(R_i(x_{i_1}, \dots, x_{i_{k_i}}))$, $i \in I$. ■

Theorem 4.7. *3-SAT implies BPI, in ZF.*

Proof. As mentioned above, $\mathcal{R}_{sat} \prec \mathcal{R}_{3-sat}$; thus, the previous theorem yields, $compact(\mathcal{R}_{3-sat}) \rightarrow compact(\mathcal{R}_{sat})$, or, equivalently, 3-SAT \rightarrow SAT. However, BPI is well-known to be equivalent to the compactness of propositional logic. Since each propositional formula is equivalent to one in conjunctive normal form, BPI is equivalent to SAT; hence 3-SAT \rightarrow BPI. ■

References

- [1] Abian, A., *On the solvability of infinite systems of Boolean polynomial equations*, Colloq. Math., 21 (1970), 27-30.
- [2] Brown, C. and Purdom, P., *An empirical comparison of backtracking algorithms*, IEEE Trans. Pattern Anal. Machine Intell., 4 (1982), 309-316.
- [3] Cowen, R., *Some combinatorial theorems equivalent to the prime ideal theorem*, Proc. Amer. Math. Soc., 41 (1973), 268-273.
- [4] Cowen, R., *Two hypergraph theorems equivalent to BPI*, Notre Dame J. Formal Logic, 31(1990), 232-240.

- [5] Garey, M. and Johnson, D., *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman, San Francisco, 1979.
- [6] Howard, P. and Rubin, J., *Consequences of the Axiom of Choice*, American Math. Society, Providence, 1998
- [7] Läuchli, H., *Coloring infinite graphs and the Boolean prime ideal theorem*, Israel J. Math., 9 (1971), 422-429.
- [8] Purdom, P., *Solving satisfiability with less searching*, IEEE Trans. Pattern Anal. Machine Intell., 6 (1984), 510-513.
- [9] West, *Introduction to Graph Theory*, Prentice Hall, 1996.
- [10] Wojtylak, P., *2-SAT is not equivalent to Boolean Prime Ideal Theorem*, in *Logic at Work, Essays Dedicated to the Memory of Helen Rasiowa*, ed. Ewa Orłowska, Springer Verlag, 1998, pp. 580-583.

Department of Mathematics,
Queens College,
CUNY
Flushing, NY 11367, USA
e-mail: cowen@qcvaqa.acc.qc.edu