Complexity of Partial Satisfaction II

Karl Lieberherr[*]

Princeton University

Ernst Specker

Swiss Federal Institute of Technology
Zurich

Technical Report #293
January 1982

ABSTRACT

What is easy and where does it get hard? We give a sharp answer to this question for a natural class $F(p,q)$ of satisfiability problems by determining algebraic numbers $\tau_{p,q} (0 < \tau_{p,q} < 1)$ which separate NP-complete from polynomial problems. Namely it is shown that the fraction $\tau_{p,q}$ of the clauses of a formula in $F(p,q)$ can be satisfied in polynomial time. However, the set of formulas in $F(p,q)$ which have an assignment satisfying the fraction $\tau'(\tau' > \tau_{p,q}, \tau'$ rational) of the clauses is NP-complete. We show a similar result for the exact satisfiability problem.

Contents

* Supported by National Science Foundation grant MCS80-04490 and Forschungsinstitut fuer Mathematik, ETH Zurich. Author's addresses: Karl Lieberherr, Princeton University, Dept. of EECS, Princeton, NJ 08544. Ernst Specker, Dept. of Mathematics, ETH-Zentrum CH-8092 Zurich, Switzerland.

1. Introduction

We continue our study of the (generalized) satisfiability problem [Lieberherr/Specker (1981), Lieberherr (1982)].

The paper os organized as follows. In section 2 we prove an auxiliary result which simplifies the problem of computing τ_ψ for a given set ψ of relations. In sections 3 and 4 we compute τ_ψ for some sets of relations. In section 5 we partially solve a problem which was left open in [Lieberherr/Specker (1981)].

2. Reduction to a continuous min-max-problem

In the following we assume that the reader is familiar with [Lieberherr (1982)]. Let $\psi = \{R_1, R_2, \ldots, R_m\}$ be a finite set of relations. Let S be a ψ-formula containing relation $R_i (1 \leqslant i \leqslant m)$ for the fraction t_{R_i} of the clauses. It follows from the methods given in [Lieberherr/Specker (1981)] that (assuming that $\sum_{i=1}^{m} t_{R_i} = 1$, $t_{R_i} \geqslant 0$ $(1 \leqslant i \leqslant m)$)

$$\tau_\psi = \lim_{n \to \infty} \tau_{n,\psi} ,$$

$$\tau_{n,\psi} = \min_{\substack{t_{R_i} (1 \leqslant i \leqslant m) \\ rational}} \; \max_{\substack{0 \leqslant k \leqslant n \\ integer}} \; mean_k^n(S)$$

$$mean_k^n(S) = \sum_{\substack{all\ relations \\ R \in S}} t_R \; SAT_k^n(R)$$

$$SAT_k^n(R) = \sum_{s=0}^{r(R)} q_s(R) \frac{(k)_s (n-k)_{r-s}}{(n)_r}$$

where

t_R	is the fraction of clauses containing relation R
$r(R)$	is the rank of R
$q_s(R)$	is the number of satisfying rows in the truth table of R which contain s ones
$(\alpha)_\beta$	$\frac{\alpha!}{(\alpha-\beta)!}$, where α,β are positive integers, $\alpha \geqslant \beta$.

Let

$$\tau_\psi' = \min_{\substack{t_{R_i}(1 \leqslant i \leqslant m) \\ real \\ \sum_{i=1}^{m} t_{R_i} = 1}} \ \max_{\substack{0 \leqslant x \leqslant 1 \\ real}} appmean_x(S)$$

$$appmean_x(S) = \sum_{\substack{all\ relations \\ R \in S}} t_R\, appSAT_x(R)$$

$$appSAT_x(R) = \sum_{s=0}^{r(R)} q_s(R)\, x^s (1-x)^{r-s} .$$

Theorem 1

$$\tau_\psi = \tau_\psi'$$

Lemma 1

Let ψ be a finite set of relations and let S be a ψ-formula.

(1) For any positive integers $n', n, n' > n$: $\tau_{n,\psi} \geqslant \tau_{n',\psi}$

(2) $$\lim_{j \to \infty} mean_{jk}^{jn}(S) = appmean_{\frac{k}{n}}(S)$$

(3) For all real $x\,(0 \leqslant x \leqslant 1)$:

$$\lim_{j \to \infty} mean_{\lceil jnx \rceil}^{jn}(S) = appmean_x(S)$$

(4)

$$\lim_{j \to \infty} \max_{\substack{0 \leqslant k \leqslant jn \\ integer}} mean_k^{jn}(S) = \max_{\substack{0 \leqslant x \leqslant 1 \\ real}} appmean_x(S)$$

(5) For any positive integer $j \geqslant 1$:

$$\max_{\substack{0 \leqslant x \leqslant 1 \\ real}} appmean_x(S) \leqslant \max_{\substack{0 \leqslant k \leqslant jn \\ integer}} mean_k^{jn}(S)$$

Proof

(1)

Assume that in any ψ-formula with n' variables the fraction $\tau_{n',\psi}$ of the clauses can be satisfied. By adding dummy variables we can transform any ψ-formula with n variables $(n < n')$ into a ψ-formula with n' variables. □

(2)

We have to show that

$$\lim_{j \to \infty} \frac{(jk)_s (j(n-k))_{r-s}}{(jn)_r} = \left(\frac{k}{n}\right)^s \left(1-\frac{k}{n}\right)^{r-s} .$$

This follows from

$$\lim_{j \to \infty} \frac{(jk)(jk-1)\cdots(jk-s+1)}{(jn)(jn-1)\cdots(jn-s+1)} = \left(\frac{k}{n}\right)^s$$

and

$$\lim_{j \to \infty} \frac{(j(n-k))(j(n-k)-1)\cdots(j(n-k)-r+s-1)}{(jn-s)(jn-s-1) \quad (jn-r+1)} = \left(1 - \frac{k}{n}\right)^{r-s} .$$

□

(3)

Follows from

$$\lim_{j \to \infty} \frac{\lfloor jnx \rceil}{jn} = x$$

and (2). □

(4)

Let x_{max} be defined by

$$\sup_{\substack{0 \leqslant x \leqslant 1 \\ real}} appmean_x(S) = appmean_{x_{max}}(S) .$$

Then by (3)

$$\lim_{j\to\infty} \max_{\substack{0 \leq k \leq jn \\ integer}} mean_k^{jn}(S) \geqslant \lim_{j\to\infty} mean^{jn}_{\lceil jnx_{max} \rceil}(S)$$

$$= appmean_{x_{max}}(S)$$

□

On the other hand by (2)

$$\lim_{j\to\infty} \max_{\substack{0 \leq k \leq jn \\ integer}} mean_k^{jn}(S) \leqslant \lim \max_{\substack{0 \leq k \leq jn \\ real}} mean_k^{jn}$$

$$= \max_{\substack{0 \leq x \leq 1 \\ real}} appmean_x(S).$$

□

(5)

It follows from the proof of (1) that

$$\max_{\substack{0 \leq k \leq jn \\ integer}} mean_k^{jn}(S) \geqslant \max_{\substack{0 \leq k \leq jn \\ integer}} mean_k^{jn}(S)$$

if $j \leqslant j'$. Now the claim follows from (4). □

Proof of Theorem

Lemma 1(5) implies that $\tau_\psi' \leqslant \tau_\psi$ and Lemma 1(4) implies that $\tau_\psi' \geqslant \tau_\psi$. □

3. Exact Satisfiability

Let R_j be the relation of rank r which holds if exactly j of the r variables are true.

Theorem 2

Let $\psi = \{R_0, R_1, \ldots, R_r\}$. Then

(i) In any ψ-formula the fraction $\frac{1}{r+1}$ of the clauses can be satisfied.

(ii) There is a polynomial algorithm *MAXMEAN* which finds an assignment satisfying at least the fraction $\frac{1}{r+1}$ of the clauses in a ψ-formula.

(iii) For any rational $\tau' > \frac{1}{r+1}$ the set of ψ-formulas having an assignment satisfying at least the fraction τ' of the clauses is NP-complete.

Proof of (i).

Since $q_s(R_j) = 0$ if $s \neq j$ and $q_j(R_j) = \binom{r}{j}$ we have

$$appmean_x(S) = \sum_{j=0}^{r} t_j \binom{r}{j} x^j (1-x)^{r-j} \quad .$$

By Theorem 1 it is sufficient to show that

$$\frac{1}{r+1} = \min_{\substack{t_j (0 \leq j \leq r) \\ real \\ \sum_{j=0}^{r} t_j = 1}} \; \max_{\substack{0 \leq x \leq 1 \\ real}} appmean_x(S).$$

ψ has the property that it is not necessary to choose the maximal x in order to satisfy τ. Therefore we perform an averaging process in the following lemma.

Lemma 2

$$\int_0^1 x^j (1-x)^{r-j} dx = \frac{1}{\binom{r}{j}(r+1)}$$

Proof

Let

$$f_{j,r} = \int_0^1 x^j (1-x)^{r-j} dx$$

We show the Lemma by induction.

$$f_{0,r} = \int_0^1 (1-x)^r\,dx = -\frac{1}{r+1}(1-x)^{r+1}\Big|_0^1 = \frac{1}{r+1}$$

For the induction step we use partial integration

$$u = \frac{1}{j+1}x^{j+1}$$

$$u' = x^j$$

$$v = (1-x)^{r-j}$$

$$v' = -(r-j)(1-x)^{r-j-1}$$

$$f_{j,r} = \int_0^1 u'v\;dx = u\cdot v\,\Big|_0^1 - \int_0^1 uv'\,dx$$

$$= \frac{r-j}{j+1} f_{j+1,r}$$

Hence,

$$f_{j+1,r} = \frac{j+1}{r-j} f_{j,r}\,(0 \leqslant j < r)$$

and therefore inductively

$$f_{j+1,r} = \frac{j+1}{r-j} f_{j,r} = \frac{j+1}{r-j}\;\frac{j\cdot(j-1)\cdots 1}{r(r-1)\cdots(r-j+1)}\cdot\frac{1}{r+1} = \frac{1}{\binom{r}{j+1}}\;.$$

Lemma 3

Let

$$\overline{t} = (t_0, t_1, \ldots, t_r)\,(\sum_{j=0}^{r} t_i = 1)$$

and let

$$appmean_x\left(\overline{t}\right) = \sum_{j=0}^{r} t_j \binom{r}{j} x^j (1-x)^{r-j}\;.$$

Then there is $x_0 (0 \leqslant x_0 \leqslant 1)$ such that

$$appmean_{x_0}\left(\overline{t}\right) = \frac{1}{r+1}\;.$$

Proof

Consider

$$\int_0^1 appmean_x\left(\vec{t}\right)dx$$

$$= \sum_{j=0}^{r} t_j \binom{r}{j} \int_0^1 x^j (1-x)^{r-j}\, dx$$

$$= \sum_{j=0}^{r} t_j \binom{r}{j} \frac{1}{\binom{r}{j}(r+1)} = \frac{1}{r+1} \ .$$

The claim follows from the mean value theorem of calculus.

Lemma 4

$$\min_{\vec{t}} \max_{0 \leqslant x \leqslant 1} appmean_x\left(\vec{t}\right) = \max_{\vec{t}} \min_{0 \leqslant x \leqslant 1} appmean_x\left(\vec{t}\right) = \frac{1}{r+1}$$

Proof

Let

$$\vec{b} = \left(\frac{1}{r+1}, \frac{1}{r+1}, \ldots, \frac{1}{r+1}\right) (r+1 \textit{ dimensional}).$$

Then

$$appmean_x\left(\vec{b}\right) = \frac{1}{r+1} \sum_{j=0}^{r} \binom{r}{j} x^j (1-x)^{r-j}$$

$$= \frac{1}{r+1}$$

Therefore

$$\min_{\vec{t}} \max_{0 \leqslant x \leqslant 1} appmean_x\left(\vec{t}\right) \leqslant \frac{1}{r+1} \ ,$$

since for $\vec{t} = \vec{b}$ only the fraction $\frac{1}{r+1}$ can be satisfied (independent of x).

Also

$$\max_{\vec{t}} \min_{x} appmean_x\left(\vec{t}\right) \geq \frac{1}{r+1} ,$$

since for $\vec{t} = \vec{b}$ the minimal x satisfies the fraction $\frac{1}{r+1}$. On the other hand for any $\vec{t}$ there is an x_0 such that

$$appmean_{x_0}\left(\vec{t}\right) = \frac{1}{r+1} .$$

Therefore

$$\min_{\vec{t}} \max_{0 \leq x \leq 1} appmean_x\left(\vec{t}\right) \geq \frac{1}{r+1}$$

and

$$\max_{\vec{t}} \min_{0 \leq x \leq 1} appmean_x\left(\vec{t}\right) \leq \frac{1}{r+1} .\square$$

Proof of (ii) and (iii).

Algorithm *MAXMEAN* in [Lieberherr (1982)] guarantees to satisfy the fraction $\frac{1}{r+1}$ in polynomial time. It follows from a general result in [Schaefer (1978)] that the ψ-satisfiability problem is NP-complete (for the ψ under discussion). Then (iii) follows from Theorem 1.2 in [Lieberherr (1982)]. □

4. Satisfiability

Let $F(p,q)$ be the following class of propositional formulas in conjunctive normal form: Each clause in a formula in $F(p,q)$ contains at least p positive or q negative literals $(p,q \geq 1)$.

Let α be the solution of $(1-x)^p = x^q$ in $(0,1)$ and let $\tau_{p,q} = 1 - \alpha^q$.

Theorem 3

(i) In any formula in $F(p,q)$ the fraction $\tau_{p,q}$ of the clauses can be satisfied.

(ii) There is a polynomial algorithm MAXMEAN which finds an assignment satisfying at least the fraction $\tau_{p,q}$ of the clauses in a formula in $F(p,q)$.

(iii) For any rational $\tau' > \tau_{p,q}$ the set of formulas in $F(p,q)$ having an assignment satisfying at least the fraction τ' of the clauses is NP-complete.

This theorem and its proof extend the results and methods given in [Lieberherr/Specker (1981)]. The proof of Theorem 3(i) is given by a sequence of simplifying reductions. The result of each reduction is given as a Proposition j . The corresponding Lemma j claims that Proposition j implies the previous proposition (in the first step: Theorem 3(i))

Theorem 3(ii) is a special case of a general result proven in [Lieberherr (1981)]. The proof of Theorem 3(iii) is based on a result by [Schaefer (1978)] and the technique given in [Lieberherr/Specker (1981)].

Simplifying Reductions

Proposition 1

For all integers $n > \min(p,q)$ and for all positive integers t_1, t_2 there is an integer k $(0 \leqslant k \leqslant n)$ such that $g_{EXACT}(n,k,t_1,t_2) =$

$$\frac{\frac{t_1}{(n)_p}(n-k)_p + \frac{t_2}{(n)_q}(k)_q}{t_1 + t_2} \leqslant 1 - \tau_{p,q}$$

Lemma 5.1

Proposition 1 ==> Theorem 3(i)

Proof:

Using the techniques given in [Lieberherr/Specker (1981)] it is easy to show that the class $F(p,q)$ can be reduced to $F'(p,q) =$ { formulas having only clauses containing either exactly p positive literals or exactly q negative literals}. Furthermore, it is sufficient to consider only symmetric formulas in $F'(p,q)$.

Let S be a symmetric formula in $F'(p,q)$ which contains t_1 clauses of the form $A_1 v A_2 v ... v A_p$ and t_2 clauses of the form $\overline{A_1} v \overline{A_2} v ... v \overline{A_q}$. Then the fraction of unsatisfied clauses if k variables are set to 1 is given by $g_{EXACT}(n,k,t_1,t_2)$. □

Note that the numerator of $g_{EXACT}(n,k,t_1,t_2)$ is the expected number of unsatisfied clauses among all assignments which set k variables to 1. It is denoted by $mean_k'(S)$ for a given formula S.

First we give an outline of the proof for Proposition 1.

Outline:

Let $x = \frac{k}{n}$ and substitute r^s for any expression of the form $(r)_s$ in $g_{EXACT}(n,k,t_1,t_2)$. The resulting expression for the fraction of *unsatisfied* clauses is

$$g_{APPROX}(x,t_1,t_2) = \frac{t_1(1-x)^p + t_2x^q}{t_1 + t_2} \quad .$$

Since for all positive integers r,k,n $(k \leqslant n)$

$$\frac{(k)_r}{(n)_r} \leqslant \left(\frac{k}{n}\right)^r \quad ,$$

the inequality

$$g_{EXACT}(n,k,t_1,t_2) \leqslant g_{APPROX}(\frac{k}{n},t_1,t_2)$$

holds. Therefore it is sufficient to show that for all n and all positive integers t_1,t_2 there is an integer k such that

$$g_{APPROX}(\frac{k}{n},t_1,t_2) \leqslant 1 - \tau_{p,q}.$$

W.l.o.g. we set $t_2 = 1$, since g_{APPROX} is homogeneous in t_1,t_2.

Take the derivative of g_{APPROX} with respect to x , set it to zero and solve for t_1:

$$t_1 = \frac{q}{p} \; \frac{x^{q-1}}{(1-x)^{p-1}}$$

Substitute for t_1 in g_{APPROX} :

$$g_{APP}(x) = \frac{q \cdot x^{q-1}(1-x)^p + p \cdot x^q (1-x)^{p-1}}{q \cdot x^{q-1} + p(1-x)^{p-1}} \quad .$$

h_2 has the following intuitive meaning. Consider a formula S in $F'(p,q)$ with n variables and assume that

$$\min_{0 \leqslant k \leqslant n} mean_k'(S) = mean_{k_{\min}}'(S) \; .$$

In any such a formula S at most the fraction $g_{APP}(\frac{k_{\min}}{n})$ of the clauses can be unsatisfied. This holds since the second derivative of h_1 with respect to x is positive for any x, $0 \leqslant x \leqslant 1$, if $p \geqslant 1$ or $q \geqslant 1$ and $t_1 \neq 0$ and $t_2 \neq 0$. Therefore it is sufficient to show that for all positive integers and all real x $(0 \leqslant x \leqslant 1)$

$$g_{APP}(x) \leqslant 1 - \tau_{p,q}.$$

Compute the extremal points of g_{APP} with respect to x in (0,1). There is only one which is given by the solution of $(1-x)^p = x^q$.

Substituting x^q for $(1-x)^p$ in g_{EXACT} yields

$$g_{EXACT} = 1 - x^q.$$

Therefore the fraction $\tau_{p,q} = 1 - \alpha^q$ can be satisfied in any formula in $F'(p,q)$.

The following simple heuristic method, which was also observed by John Scranton, gives the correct result.

Choose x such that the fraction of satisfied clauses is independent of t_1, t_2. The resulting condition for x is

$$(1 - x)^p = x^q.$$

For such an x, the fraction of satisfied clauses is independent (in the limit) of the formula we consider and it is $\tau_{p,q}$.

Now we continue with the proof of Proposition 1.

Proposition 2

For all integers $n > \min(p,q)$ and all positive integers t_1, t_2 there is an integer $k\,(0 \leqslant k \leqslant n)$ such that

$$g_{APPROX}(x, t_1, t_2) = \frac{t_1(1-x)^p + t_2 x^q}{t_1 + t_2} \leqslant 1 - \tau_{p,q}$$

Lemma 5.2

Proposition 2 ==> Proposition 1

Proof:

Note that for all positive integers $r, k, n\,(k \leqslant n)$ since

$$\frac{(k)_r}{(n)_r} \leqslant \left(\frac{k}{n}\right)^r$$

$$(1-\frac{1}{k})(1-\frac{2}{k})\cdots(1-\frac{r-1}{k}) \leqslant (1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{r-1}{n}) .$$

If we let $x = \frac{k}{n}$ and replace $\frac{(n-k)_p}{(n)_p}$ by $(1-x)^p$ and $\frac{(k)_q}{(n)_q}$ by x^q we increase g_{EXACT}. Therefore $g_{EXACT} \leqslant g_{APPROX}$ which proves Lemma 2. □

Proposition 3

For all real x $(0 \leqslant x \leqslant 1)$

$$g_{APP} = \frac{x^{q-1}(1-x)^{p-1}[q(1-x)+px]}{q \cdot x^{q-1} + p(1-x)^{p-1}} \leqslant \alpha^q$$

Lemma 5.3

Proposition 3 ==> Proposition 2

Proof

W.l.o.g. we set $t_2 = 1$ since g_{APPROX} is homogeneous in t_1, t_2. Take the derivative of g_{APPROX} with respect to x, set it to zero and solve for t_1:

$$t_1 = \frac{q}{p} \frac{x^{q-1}}{(1-x)^{p-1}}$$

If we substitute for t_1 in g_{APPROX} we get g_{APP}. Note that the second derivative of g_{APPROX} with respect to x is

$$t_1 \cdot p \cdot (p-1)(1-x)^{p-2} + t_2 \cdot q \cdot (q-1)x^{q-2},$$

which is positive for any $x\,(0 \leqslant x \leqslant 1)$, if $p \geqslant 1$ or $q \geqslant 1$ and $t_1 \neq 0$ and $t_2 \neq 0$.

Proof of Proposition 3

We show first that the derivative of $g_{APP}(x)$ is zero in $(0,1)$ iff x satisfies $(1-x)^p = x^q$.

Let $A = x^q, B = (1-x)^p$. Then

$$g_{APP}(x) = \frac{A'B - AB'}{A' - B'}$$

The numerator of the derivative of $g_{APP}(x)$ is

$$(A''B' - A'B'')\ (A - B)$$

The first factor

$$A''B' - A'B''$$

$$= -pq(q-1)x^{q-2}(1-x)^{p-1} - q\cdot p(p-1)x^{q-1}(1-x)^{p-2}$$

$$= x^{q-2}(1-x)^{p-2}(-(q-1)(1-x) - (p-1)x)$$

has no zeros in $(0,1)$.

Since $(1-x)^p = x^q$ has only one solution α in $(0,1)$ the rational function $g_{APP}(x)$ has one extremal point in $(0,1)$ with value $g_{APP}(\alpha) = \alpha^q$. Since $g_{APP}(0) = g_{APP}(1) = 0$ the function $g_{APP}(x)$ is maximal for $x = \alpha$. □

The proof of Proposition 3 uses differentiation. We give now a different proof which does not use differentiation and which provides further insight into the problem.

Proposition 4

For all real x,β $(0 \leqslant x,\beta \leqslant 1)$

$$g_2(x,\beta) = x^q\left[(1-x)^p(px-qx+q) + (1-\beta)^p q(x-1)\right] - (1-x)^p\beta^q\cdot p\cdot x \leqslant 0$$

Lemma 5.4

Proposition 4 ==> Proposition 3

Proof

Multiply both sides of $g_{APP}(x) \leqslant \alpha^q$ by the denominator of $g_{APP}(x)$ and shift all terms to the left of the inequality sign. The resulting inequality is $g_2(x,\beta) \leqslant 0$ if we make liberal use of $(1-\alpha)^p = \alpha^q$ (a crucial point) and if we substitute β for α.

Proof of Proposition 4

So far a proof of Proposition 4 was obtained only for special cases.

1) $p=1, q \geqslant 1$.

Note that

$$g_2(x,\beta) = (x-\beta)^2 \sum_{i=1}^{q-1} x^{q-i-1}(i-q)\beta^{i-1}$$

Hence $g_2(x,\beta)$ is non-positive for $0 \leqslant x,\beta \leqslant 1$.

Example: $(p=1)$

$g_2(x,\beta)$ is proportional to (the deleted factor has a positive sign)

$-(x-\beta)^2$for $q=2$.

$-(x-\beta)^2(2x+\beta)$ for $q=3$.

$-(x-\beta)^2(3x^2+2x\alpha+\alpha^2)$ for $q=3$.

II) $p=2$.

$g_2(x,\beta)$ is proportional to (the deleted factor has a positive sign)

$(x-\beta)^2(x(x-4) + 2\beta(x-1))$ for $q=3$.

$(x-\beta)^2(x^2(x-3) + 2\beta x(x-1) + \beta^2(x-1))$ for $q=4$.

$(x-\beta)^2(x^3(3x-8) + 6\beta x^2(x-1) + 4\beta^2 x(x-1) + 2\beta^3(x-1))$ for $q=5$.

Unfortunately this technique does not generalise for $p \geqslant 3$ but it is conjectured that Proposition 5 holds in general.

The formulas obtained by the alternate proof method have interesting applications.

Theorem 4:

Let S be a formula in $F'(1,q)(q>1)$ for which

$$\max_{0 \leqslant k \leqslant n} mean_k(S) = mean_0(S).$$

Then assignment $J_{ALL\ 0}$, which assigns false to all variables satisfies all clauses. In general, if $\max_{0 \leqslant k \leqslant n} mean_k(S) = mean_{k'}(S)$ then the assignment which assigns true to k' variables satisfies at least the fraction

$$1 - \alpha^q \frac{(\frac{k'}{n-\alpha})^2 \sum_{i=1}^{q-1} (i-q)\alpha^{i-1}(\frac{k'}{n})^{q-i-1}}{q\cdot(\frac{k'}{n})^{q-1} + 1}$$

of the clauses.

Proof

Consider $g_{APP}(x) - \alpha^q$ for $p=1$ (after multiplying with $q\cdot x^{q-1} + 1$):

$$qx^{q-1}(1-x) + x^q - \alpha^q \cdot (q\cdot x^{q-1}+1) = (x-\alpha)^2 \sum_{i=1}^{q-1} x^{q-i-1}(i-q)\alpha^{i-1}.$$

Let

$$h(x) = \frac{(x-\alpha)^2 \sum_{i=1}^{q-1} x^{q-i-1}(i-q)\alpha^{i-1}}{q\cdot x^{q-1}+1} .$$

Now,

$$h(0) = \alpha^2(-1\cdot\alpha^{q-2}) = -\alpha^q .$$

□

Proof of Theorem 3(ii)

Algorithm MAXMEAN in [Lieberherr (1981)] guarantees to satisfy the fraction $\tau_{p,q}$ in polynomial time. □

Proof of Theorem 3(iii)

The fact that the satisfiability problem for formulas in $F(p,q)$ is NP-complete follows from a general result of [Schaefer (1978)]. Then the proof can be adapted from [Lieberherr/Specker (1981)].

Extensions

The technique used to prove Theorem 3(i) is suitable to determine τ_ψ for other sets ψ which contain only two relations. The fraction of satisfied clauses in a symmetric formula which contains t_1 clauses with the first relation and t_2 clauses with the second relation is given by (in approximated form)

$$h_1(x,t_1,t_2) = \frac{t_1R_1(x) + t_2R_2(x)}{t_1 + t_2} ,$$

where R_1 and R_2 are polynomials which depend on the two relations.

W.l.o.g. $t_2 = 1$. If we take the derivative of h_1 with respect to x and solve for t_1 we get

$$t_1 = \frac{-R_2'(x)}{R_1'(x)} .$$

Substituting in h_1 we get

$$h_2(x) = \frac{R_1'(x)\cdot R_2(x) - R_1(x)R_2'(x)}{R_1'(x) - R_2'(x)} .$$

The numerator of the derivative of $h_2(x)$ is given by

$$(R_1(x) - R_2(x))(R_1''(x)R_2'(x) - R_1'(x)R_2''(x)) .$$

If the second factor has no zeros in (0,1) then the fraction $R_1(\alpha)$ can always be satisfied, where α is the solution of $R_1(x) = R_2(x)$ in (0,1) which is the global minimum of h_2.

5. Partial Solution of the 3-Satisfiability Problem

In [Lieberherr/Specker (1981)] the following problem was left open. A formula S of the propositional calculus in conjunctive normal form is said to be *3-satisfiable*, if any triple of clauses is satisfiable. Determine the fraction τ_3 of the clauses which can always be satisfied in a 3-satisfiable formula. We show in the following that $\tau_3 \geqslant 2/3$. The problem with 3-satisfiable formulas is that they are not closed under symmetrization. If we take a 3-satisfiable formula S and symmetrize it with the full permutation group then the symmetrized formula is in general not 3-satisfiable.

To show that $\tau_3 \geqslant 2/3$ we construct a class RED_1 of formulas so that

1. RED_1 contains all 3-satisfiable formulas (but some are not 3-satisfiable)

2. in any formula in RED_1 at least the fraction 2/3 of the clauses can be satisfied.

Consider any 3-satisfiable formula S. Without loss of generality we assume that clauses of length 1 only contain positive literals (this can be enforced by renamings). Now we partition the variables into two classes. The first class contains only variables which occur in clauses of length 1. The second class contains all other variables. A clause is said to be type T_{ij}^q if its j variables are in class q and i of them are positive. A clause is said to be of type $T_{i_1 j_1 i_2 j_2}^{qr}$ if it contains j_1 variables of class q and j_2 variables of class r and if i_1 of the j_1 variables are positive and i_2 of the j_2 variables are positive.

Definition

RED_1 is the following subset of cnfs: The variables are partitioned into 2 classes and only the following clause types occur:

$$T_{11}^1, T_{03}^1, T_{01\,11}^{12}, T_{01\,01}^{12}, T_{02}^2, T_{12}^2, T_{22}^2 \ .$$

This definition is of interest since for proving that $\tau_3 \geqslant 2/3$ it is sufficient to minimize among the formulas in RED_1.

Theorem 5:

(i) In any 3-satisfiable cnf at least the fraction 2/3 of the clauses can be satisfied. (ii) There is a polynomial algorithm to find such an assignment.

Proposition 2.1

In any cnf in RED1 at least the fraction 2/3 of the clauses can be satisfied.

Lemma 2.1

Proposition 2.1 ==> Theorem 2(i)

Proof:

Any 3-satisfiable cnf is easily reduced to a formula in RED_1 by deleting literals. Deleting literals makes a formula harder for satisfying many clauses. □

Definition:

Let RED_2 be the subset of cnfs of RED_1 which do not contain clauses with types T^2_{02}, T^2_{12} and T^2_{22}.

Proposition 2.2

In any cnf in RED2 at least the fraction 2/3 of the clauses can be satisfied.

Lemma 2.2

Proposition 2.2 ==> Proposition 2.1

Proof:

In a cnf containing clauses of exactly length 2 at least the fraction 3/4 of the clauses can be satisfied (a random assignment satisfies 3/4). Therefore deleting clauses of the above three types does not make it easier to satisfy many clauses. □

We prove now Proposition 2.2 by a sequence of further reductions. Let S be a formula in RED_2 which contains t_1 clauses of type T^1_{11}, t_2 clauses of type $T^{12}_{01\,11}, t_3$ clauses of type T^1_{03} and t_4 clauses of type $T^{12}_{01\,11}$. The worst-case formulas (regarding the fraction of satisfiable clauses) in RED_2 are those which are symmetric in the A-variables and B-variables. Among those formulas the formulas with $t_2 = t_4$ are hardest. In a formula in RED_2 with $t_2 = t_4$ the fraction

$$1 - \frac{\frac{t_1}{n}(n-k) + \frac{t_2}{n}k + \frac{t_3}{\binom{n}{3}}\binom{k}{3}}{t_1 + 2t_2 + t_3}$$

of the clauses are satisfied if k of the n A-variables are set to 1. Therefore, we have to show

Proposition 2.3

For all integers n and for all positive integers t_1, t_2, t_3 there is an integer k $(0 \leqslant k \leqslant n)$ such that

$$\frac{\frac{t_1}{n}(n-k) + \frac{t_2}{n}k + \frac{t_3}{(n)_3}(k)_3}{t_1 + 2t_2 + t_3} \leqslant \frac{1}{3}.$$

Lemma 2.3

Proposition 2.3 ==> Proposition 2.2

Proof:

Given above. □

Proposition 2.4

For all integers n and for all positive integers t_1, t_2, t_3 there is an integer $k\,(0 \leqslant k \leqslant n)$ such that

$$w_1 = \frac{\frac{t_1}{n}(n-k) + \frac{t_2}{n}k + \frac{t_3}{n^3}k^3}{t_1 + 2t_2 + t_3} \leqslant \frac{1}{3} .$$

Lemma 2.4

Proposition 2.4 ==> Proposition 2.3

Proof:

Observe that $\frac{(k)_r}{(n)_r} \leqslant \frac{k^r}{n^r}$ if $k \leqslant n$. □

Proposition 2.5

For all x $(0 \leqslant x \leqslant 1)$ and all positive integers t_3

$$w_2 = \frac{1 - 2t_3x^3}{3 + t_3(1-6x^2)} \leqslant \frac{1}{3} .$$

Lemma 2.5

Proposition 2.5 ==> Proposition 2.4

Proof:

W.l.o.g. let $t_1 = 1$ and substitute x for $\frac{k}{n}$ in w_1. Take the derivative of w_1 with respect to x, set it to zero and solve for t_2:

$$t_2 = 1 - 3t_3x^2.$$

By substituting $1 - 3t_3x^2$ for t_2 in w_1 we get w_2. □

Proposition 2.5 is easily proven directly by case analysis. □

6. Conclusions and open problems

The analysis of algorithm $\psi-MAXMEAN$ reduces to a minimax problem of the following form: Let

$$\vec{t} = (t_1, \ldots, t_m),\ 0 \leq t_i \leq 1,\ t_i\ real,\ \sum_{i=0}^{m} t_i = 1.$$

Let $\vec{Q} = (Q_1(x),...,Q_m(x))$ be a vector of m polynomials with integer coefficients. The minimax problem consists of computing

$$\tau_{\vec{Q}} = \min_{\vec{t}} \max_{0 \leq x \leq 1} \sum_{i=1}^{m} t_i Q_i(x) .$$

It appears that there exist no efficient methods for computing $\tau_{\vec{Q}}$, given $\vec{Q}$. Methods of classical game theory [von Neumann/Morgenstern (1947)] and their extensions [Fan (1953)] help only in special cases.

In the area of robust estimation, where related minimax problems are analyzed, no efficient methods for solving them have been developed either (see [Vastola/Poor (1981)]). All these facts suggest the conjecture that the following problem MINIMAX is NP-complete:

Input: A vector $\vec{Q}$ of polynomials with integer coefficients and a rational number r.

Question: $\tau_{\vec{Q}} \leq r$

References

Fan1953a. K. Fan, "Minimax Theorems," *Proc. Nat. Acad. Sci.* **39**, pp.42-47 (1953).

Lieberherr1981b. K. Lieberherr and E. Specker, "Complexity of partial satisfaction," *Journal of the ACM* **28**(2), pp.411-421 (1981).

Lieberherr1981a. K.J. Lieberherr, "Probabilistic combinatorial optimization," *Lecture Notes in Computer Science, 10th International Symposium on Mathematical Foundations of Computer Science '81, Strbske Pleso, Czechoslovakia*, Springer (1981).

Lieberherr1982a. K.J. Lieberherr, "Algorithmic extremal problems in combinatorial optimization," *Journal of Algorithms, in print* (1982).

Neumann1947a. J. von Neumann O. Morgenstern, *Theory of Games and Economic Behavior,* Princeton University Press, Princeton (1947).

Schaefer1978a. T. Schaefer, "The complexity of satisfiability problems," *Proc. 10th Annual ACM Symposium on Theory of Computing,* pp.216-226 (1978).

Vastola1981a. K.S. Vastola H.V. Poor, "Robust Wiener-Kolmogorov Theory," *preprint, submitted for publication* (1981).