離散アルゴリズムの高品質化・高速化の追究とその応用に関する研究

研究代表者　研　究　員　**浅野　孝夫**（中央大学理工学部情報工学科）
共同研究者　研　究　員　**今井　桂子**（中央大学理工学部情報工学科）

1　はじめに

精度保証のある近似アルゴリズムの研究は，最近近似アルゴリズムの理論の枠組みが統一化されるに従い，以前にもまして，急速に進展している。特に，Goemans-Williamson[4] の MAX CUT（最大カット問題）と MAX 2SAT に対する半正定値計画（以下 SDP と略記する）緩和に基づく方法は，従来の近似アルゴリズムの設計解析手法の枠組みを越えて，近似アルゴリズムにおけるブレークスルーをもたらした斬新で画期的な手法として今日認識されている。本文は，SDP 緩和に基づいた近似アルゴリズムとして，MAX SAT に関連する話題に焦点をあてて，我々の研究成果と最近の研究動向を述べる。

2　MAX SAT とは

MAX SAT（maximum satisfiability problem，**最大充足化問題**）は，充足可能性問題の最適化版ともいえるが，入力は論理和の集合 C と各論理和 $C \in C$ に対する非負の重み $w(C)$ の対 (C, w) で規定される。$X=\{x_1, x_2, \cdots, x_n\}$ を C の論理和に現れる変数の集合とする。すると，MAX SAT は，各変数 $x_i \in X$ に対して，真偽割当をして満たされる（1 となる）論理和の重みの総和を最大にする問題である。$\bar{x}_i=1-x_i$ であるので $C_j \in C$ は

$$C_j(x)=1-\prod_{x_i \in X_j^+}(1-x_i)\prod_{x_i \in X_j^-} x_i \qquad (1)$$

として $\boldsymbol{x}=(x_1, x_2, \cdots, x_n)$ の関数と考えることができる。ここで X_j^+ は C_j に肯定形であらわれる変数の集合で X_j^- は否定形であらわれる変数の集合である。こうして，任意の真偽割当 $\boldsymbol{x} \in \{0,1\}^n$ に対し C_j は $C_j(\boldsymbol{x})=0$ または 1 となり，真偽割当 x の値は

$$F(x)=\sum_{C_j \in C} w(C_j)C_j(\boldsymbol{x}) \qquad (2)$$

と定義される。すなわち $\boldsymbol{x}$ によって満たされる C 内の論理和の重みの和が $\boldsymbol{x}$ の値である。こうして MAX SAT は値が最大となるような $\boldsymbol{x}$ を見つける問題となる。なお，MAX SAT の入力 C の各論理和が k 個以下のリテラルからなるときは，MAX SAT は **MAX kSAT** と呼ばれている。MAX 2SAT も有名な NP- 困難（MAX-SNP 困難）問題である。MAX SAT に対する従来の近似アルゴリズムは条件付き確率法に基づく方法であるが，ここでは，最新の SDP 緩和に基づく方法を述べる。

MAX SAT		MAX kSAT	
0.7584	GW95[4]	0.878（2）	GW95[4]
0.765	AOH96	0.931（2）	FG95[3]
0.770	A97[1]	0.801（3）	TSSW
0.797 ‡	Z99[9]	0.875（3）	KZ97[7]
0.8331 ‡	AW00[2]	0.8721†（4）	HZ99[5]

図 1　MAX SAT に対する近似率。†は[5]での数値実験で示された近似率。‡は[9]で予想された近似率に基づくもの。（　）内の数字は MAX kSAT のk。

3　SDP 緩和に基づく方法

図 1 は MAX SAT および MAX kSAT に対する近似アルゴリズムの性能比較である。いずれも SDP 緩和に基づいている。以下，MAX 2SAT に対する Goemans-Williamson[4]，Feige-Goemans[3]，Zwick[8] による SDP 緩和法について説明する。いずれも

（A）問題の SDP 緩和による定式化を求める。
（B）SDP 緩和問題を解き最適解のベクトルを多項式時間で求める。
（C）ランダムベクトルを用いて各ベクトルを 0,1 に丸めて変数へのランダム真偽割当を求める。
（D）ランダム真偽割当の期待値以上の値をもつ真偽割当を求める。

から成り立っている。以下では，（A）と（C）に注目する（実際最近の SDP 緩和に基づく近似アルゴリズムはこの部分の研究が主になっている）。簡単のため $x_{n+i}=\bar{x}_i$（$x_i=\bar{x}_{n+i}$，$x_i+x_{n+i}=1$）とおく。論理和 $x_i \vee x_j$ を C_{ij} と表し，C_{ij} の重みを w_{ij} と書く。1 つのリテラルからなる論理和 x_i も $x_0 \equiv 0$ を用いて便宜的に 2 つのリテラルからなる論理和 $x_0 \vee x_i$ と書けることになる。リテラルへの真偽割当 $\boldsymbol{x}=(x_0, x_1, \cdots, x_{2n}) \in \{0,1\}^{2n+1}$ に対して論理和 $C_{ij}=x_i \vee x_j$ の値 $z_{ij}(\boldsymbol{x})$ および $\boldsymbol{x}$ の値 $F(\boldsymbol{x})$ は（1），（2），より $z_{ij}(\boldsymbol{x})=1-(1-x_i)(1-x_j)$，$F(\boldsymbol{x})=\sum_{C_{ij} \in C} w_{ij} z_{ij}(\boldsymbol{x})$ と書ける。

3.1　Goemans-Williamson の方法

Goemans-Williamson の MAX　2SAT に対するアルゴリズム[4]は最初に

$$y_0=-1, \qquad y_0y_i=1-2x_i$$

とおいて，(0, 1)- 割当 $\boldsymbol{x}=(x_0, x_1, \cdots, x_{2n}) \in \{0, 1\}^{2n+1}$ から(－1, 1)- 割当 $\boldsymbol{y}=(y_0, y_1, \cdots, y_{2n}) \in \{-1, 1\}^{2n+1}$に一対一対応させておく($x_{n+i}=1-x_i$より$y_{n+i}=-y_i$)。したがって，$x_i$は $\frac{1-y_0y_i}{2}$ となり，論理和 $C_{ij}=x_i \vee x_j$の値は$\boldsymbol{y}=(y_0, y_1, \cdots, y_{2n})$の関数として

$$\begin{aligned} z_{ij}(\boldsymbol{y}) &= 1-\frac{1+y_0y_i}{2}\frac{1+y_0y_j}{2} \\ &= \frac{3-y_0y_i-y_0y_j-y_iy_j}{4} \end{aligned}$$

と書ける。さらに，$\boldsymbol{y}$ の値は$F(\boldsymbol{y})=\sum_{c_{i,j} \in c} w_{ij} z_{ij}(\boldsymbol{y})$となり，MAX 2SAT は$F(\boldsymbol{y})=\sum_{c_{i,j} \in c} w_{ij} z_{ij}(\boldsymbol{y})$を最大にするような(－1, 1)- 割当$\boldsymbol{y}=(y_0, y_1, \cdots, y_{2n})$を求める問題となる($y_0=-1, y_{n+i}=-y_i$)。

半正定値計画問題に緩和するため，y_iに対応してノルム 1 でさらに $i \neq 0$ のときは $\boldsymbol{v}_{n+i}=-\boldsymbol{v}_i$ を満たす($2n+1$)- 次元のベクトル $\boldsymbol{v}_i \in S^{2n}$ を導入する($i=0, 1, \cdots, 2n$)。こうして，各 $y_{i_1}y_{i_2}$ をベクトルの内積 $\boldsymbol{v}_{i_1}\boldsymbol{v}_{i_2}$ に置き換え，$y_{i_1i_2}=\boldsymbol{v}_{i_1}\boldsymbol{v}_{i_2}$ とする($Y=(y_{i_1i_2})$は半正定値行列になる)。このように，Goemans-Williamson のアルゴリズムは絶対値 1 の変数 y_i を R^{2n+1} における単位球 S^{2n} 上のベクトル $\boldsymbol{v}_i$ で緩和して $2n+1$ 個のベクトル $\boldsymbol{v}_0, \boldsymbol{v}_1, \cdots, \boldsymbol{v}_{2n}$ を用いるアルゴリズムである。したがって，MAX 2SAT に対するベクトル計画緩和は，

$$\begin{aligned} \text{(GW)}\ \max\ & \sum_{C_{ij} \in C} w_{ij} z_{ij} \\ \text{s.t.}\ \ & z_{ij}=\frac{3-\boldsymbol{v}_0\boldsymbol{v}_i-\boldsymbol{v}_0\boldsymbol{v}_j-\boldsymbol{v}_i\boldsymbol{v}_j}{4} && (C_{ij} \in C) \\ & \boldsymbol{v}_i \in S^{2n} && (0 \leq i \leq 2n) \\ & \boldsymbol{v}_{n+i}\boldsymbol{v}_i=-1 && (1 \leq i \leq n) \end{aligned}$$

となる。ベクトル $\boldsymbol{v}_i$ を $\boldsymbol{v}_0=-1$，$\boldsymbol{v}_i \in \{-1, 1\}$とすれば MAX 2SAT と等価であることが確認できる。ベクトル計画問題と半正定値計画問題が等価であるので，ここでは用語を互いに区別せずに用いる。

Goemans-Williamson のアルゴリズムはまずこの問題(GW)を多項式時間で解き最適解のベクトル $\boldsymbol{v}_0, \boldsymbol{v}_1, \cdots, \boldsymbol{v}_{2n}$ を求める。この解から(－1, 1)- 割当$\boldsymbol{y}=\{y_0, y_1, \cdots, y_{2n}$を求めることを**ラウンディング**という。Goemans-Williamson のラウンディングは，原点を通るランダムな超平面で空間を 2 つの領域に分割し，$\boldsymbol{v}_i$ が $\boldsymbol{v}_0$ と同じ領域に入るとき$y_i=-1$($x_i=0$)とし，反対の領域に入るとき$y_i=1$($x_i=1$)とするものである。

このランダム真偽割当により論理和 $C_{ij}=x_i \vee x_j$の満たされる確率 p_{ij} は，ベクトル $\boldsymbol{v}_0, \boldsymbol{v}_i, \boldsymbol{v}_j$ から定まる $\boldsymbol{v}_0\boldsymbol{v}_i=\cos\ \theta_{0i}$，$\boldsymbol{v}_0\boldsymbol{v}_j=\cos\ \theta_{0j}$，$\boldsymbol{v}_i\boldsymbol{v}_j=\cos\ \theta_{ij}$ を用いると，$\frac{\theta_{0i}+\theta_{0j}+\theta_{ij}}{2}$ に等しい。論理和 $C_{ij}=x_i \vee x_j$ が 1 個のリテラル x_i からなるときは，$x_j=x_0$として，論理和 $C_{0i}=x_0 \vee x_i$の満たされる確立 p_{0i} は $\frac{\theta_{0i}}{\pi}$ になる。1 個のリテラルからなる論理和 $C_{0i}=x_0 \vee x_i$の満たされる確率 p_{0i} と最適解における値 $z_{0i}=\frac{1-\boldsymbol{v}_0\boldsymbol{v}_i}{2}$ の比 $\frac{p_{0i}}{z_{0i}}$ は少なくとも

$$\alpha_1 \equiv \min_{0 \leq \theta_{0i} \leq \pi} \frac{2\theta_{0i}}{\pi(1-\cos\theta_{0i})}$$

である。最小値を実現する θ_{0i} の値 $\theta_{0i}^{(1)}$は $\frac{\theta}{1-\cos\theta}$ の θ に関する微分が $\frac{1-\cos\theta-\theta\sin\theta}{(1-\cos\theta)^2}$ であることから唯一に $\theta_{0i}^{(1)}=2.331122$ と定まり，$\alpha_1=0.878567$ となる。一方，2 個のリテラルからなる論理和 $C_{ij}=x_i \vee x_j$の満たされる確率 p_{ij} と最適解における値 $z_{ij}=\frac{3-\boldsymbol{v}_0\mathrm{v}_\mathrm{i}-\boldsymbol{v}_0\boldsymbol{v}_j-\boldsymbol{v}_i\boldsymbol{v}_j}{4}$ の比 $\frac{p_{ij}}{z_{ij}}$ は少なくとも

$$\alpha_2 \equiv \min_{0 \leq \theta_{0i},\ \theta_{0j},\ \theta_{ij} \leq \pi} \frac{2(\theta_{0i}+\theta_{0j}+\theta_{ij})}{\pi(3-\cos\theta_{0i}-\cos\theta_{0j}-\cos\theta_{ij})}$$

となる。なお，$\alpha_2=\alpha_1$ であることも得られる。

このことから，論理和全体の期待値の最適解の値に対する比は 0.87856 以上となり，期待値以上の割当は必ず存在し，そのようなものを見つける方法も存在するので，Goemans-Williamson のアルゴリズムは 0.87856-近似アルゴリズムである。

3.2　Feige-Goemans の回転

Feige-Goemans[3] は，Goemans-Williamson のラウンディングで，最小の性能 0.87856 を実現するベクトルの組に注目し，そのような状況を避けるようにできれば，近似率の改善につながるはずと確信し，以下の手法を考案した。彼らは，ランダムな超平面でベクトルを分割する前に，各 $\boldsymbol{v}_i (i \neq 0)$を $\boldsymbol{v}_0$ と $\boldsymbol{v}_i$ で定まる平面上で，そのなす角 θ_{0i} に基づいて回転し，得られた $\boldsymbol{v}_i'(i \neq 0,\ \boldsymbol{v}_0'=\boldsymbol{v}_0)$に対して Goemans-Williamson のラウンディングを適用している。より具体的には，2 個のリテラルからなる論理和がなければ，原理的には，1 個のリテラルからなる論理和 $C_{0i}=x_0 \vee x_i$の満たされる確立 $\frac{\theta_{0i}}{\pi}$ の最適解に対する比が 1 になるように，$\boldsymbol{v}_0$ と $\boldsymbol{v}_i$ で定まる平面上で，回転後のベクトルの角度 θ_{0i}' が $\theta_{0i}'=\frac{\pi}{2}(1-\cos\theta_{0i})$を満たすようにすればよい。しかしながら，2 個のリテラルからなる論理和がある時にはこの回転でそのような論理和

$C_{ij}=x_i \vee x_j$の満たされる確率の最適解に対する比$\frac{p_{ij}}{z_{ij}}$が減少してしまうこともある。そこで回転角を定める関数$f:[0,\pi]\to[0,\pi]$をパラメータ$0\le\lambda\le 1$を用いて

$$f(\theta)=(1-\lambda)\theta+\lambda\frac{\pi}{2}(1-\cos\theta)$$

としている。こうして$\boldsymbol{v}_0$，$\boldsymbol{v}_i$がなす平面上で$\boldsymbol{v}_i$を回転して得られる$\boldsymbol{v}_i'$は，$\boldsymbol{v}_0$と$\theta_{0i}'=f(\theta_{0i})$の角度をなすようにした。直観的には，なす角が90度から離れるようにして，ランダム性を排除しているといえる（どの2つのベクトルも直交するときランダムということになるので）。また，fは$f(\pi-\theta)=\pi-f(\theta)$を満たしていることに注意されたい。この関数による$\boldsymbol{v}_i$，$\boldsymbol{v}_j$のなす角$\theta_{ij}$と$\boldsymbol{v}_i'$，$\boldsymbol{v}_j'$のなす角$\theta_{ij}'$は，$\boldsymbol{v}_0$と$\boldsymbol{v}_i$がなす平面と，$\boldsymbol{v}_0$と$\boldsymbol{v}_j$がなす平面の角度を$\alpha$とおくと，一般性を失うことなく，3次元正規直交$xyz$座標形で$\boldsymbol{v}_0=(0,0,1)$，$\boldsymbol{v}_i=(\sin\theta_{0i},0,\cos\theta_{0i})$，$\boldsymbol{v}_j=(\cos\alpha\sin\theta_{0j},\sin\alpha\sin\theta_{0j},\cos\theta_{0j})$と仮定できるので，

$$\cos\theta_{ij}'=\cos\theta_{0i}'\cos\theta_{0j}'+\cos\alpha\sin\theta_{0i}'\sin\theta_{0j}'$$

$$\cos\theta_{ij}=\cos\theta_{0i}\cos\theta_{0j}+\cos\alpha\sin\theta_{0i}\sin\theta_{0j}$$

をみたすことがいえる。彼らは$\lambda=0.806765$としてGoemans-Williamsonのラウンディングをすると，得られる真偽割当は2個のリテラルからなる論理和$x_i \vee x_j$において満たされる確率p_{ij}と最適解における値の比$\frac{p_{ij}}{z_{ij}}=\frac{2(\theta_{0i}'+\theta_{0j}'+\theta_{ij}')}{\pi(3-\cos\theta_{0i}-\cos\theta_{0j}-\cos\theta_{ij})}$は0.93109（1個のリテラルからなる論理和に対してこの比は0.976）以上になることを数値的に示し，0.93109-近似アルゴリズムが実現できることを示した。

3.3　Zwickの回転

Zwick[8]は，$f(\theta)=0\,(\theta<\frac{\pi}{2})$，$\frac{\pi}{2}\,(\theta=\frac{\pi}{2})$，$\pi\,(\theta>\frac{\pi}{2})$を用いてベクトルを回転しその後Goemans-Williamsonのラウンディングをすると，充足可能なMAX 2SATのインスタンスに対してはどの論理和も満たされることを示し，それに基づいて，回転関数を

$$f_{\epsilon^{1/3}}(\theta)=\begin{cases}0 & 0\le\theta<\frac{\pi}{2}-\epsilon^{1/3},\\ \frac{\pi}{2}+\frac{\pi}{2\epsilon^{1/3}}(\theta-\frac{\pi}{2}) & \frac{\pi}{2}-\epsilon^{1/3}\le\theta\le\frac{\pi}{2}+\epsilon^{1/3},\\ \pi & \frac{\pi}{2}+\epsilon^{1/3}<\theta\le\pi\end{cases}$$

として定義した。$Z=\sum_{c_{ij}\in c}w_{ij}z_{ij}$，$W=\sum_{c_{ij}\in c}w_{ij}$，$\epsilon=1-Z/W$で，$\epsilon$は小さいものとしている。この関数$f_{\epsilon^{1/3}}(\theta)$はほぼ充足可能なMAX 2SATのインスタンス，すなわち，$\epsilon=1-Z/W$の十分小さなインスタンスに対して良い性能を示す。

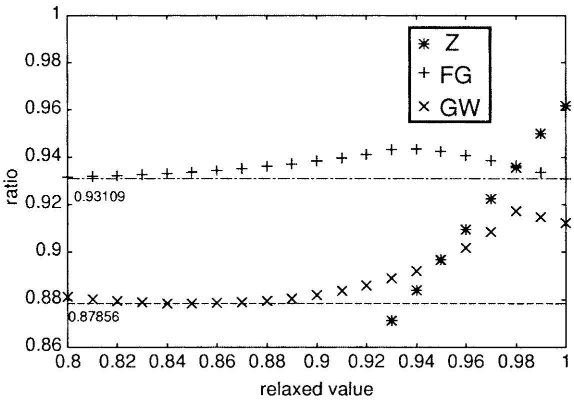

図2　z_{ij}に対する各関数の性能

3.4　実験結果

図2はこれまで説明してきたGoemans-Williamson，Feige-Goemans，ZwickのSDP緩和問題の最適解の各論理和C_{ij}の値z_{ij}と回転した後の乱数使用ラウンディングの論理和C_{ij}の満たされる確率p_{ij}の比$\frac{p_{ij}}{z_{ij}}$を表したものである。図3と図4はさらにFeige-GoemansとZwickの詳細である。これからもわかるように，Zwickの方法では精度保証は得られない。

岩間-浅野[6]はこれらのアルゴリズムの実際的性能評価をするための計算機実験を行なっている。詳細はそちらを参照されたいが，ほぼすべての入力に対して同様の傾向が観察された。より具体的例を挙げよう。50変数，450個の論理和からなる入力で，$\epsilon=0.12348$のとき各論理和C_{ij}の最適解における値z_{ij}を$0\le z_{ij}\le 0.01$，$0.01<z_{ij}\le 0.02$，…，$0.99<z_{ij}\le 1$となるように0.01きざみで分類して，その個数をヒストグラムとして表したグラフが図5である。最も多かったのは$0.99<z_{ij}\le 1$を満たす論理和であり，450個の論理和中186個の論理和が存在した。また，z_{ij}の値が1に近い$z_{ij}>1-\epsilon=0.87652$となる論理和は全論理和の72%程度であった。

このように，SDP緩和によって得られた各z_{ij}の値の分布を図5より観察すると，大半のz_{ij}は1に近い値をとっている。図4よりZwickの関数はz_{ij}が大きくて1に近いときに良い性能を示すのに対して，Feige-Goemansの関数は，図3より，z_{ij}が1に近いときは，最悪に近いときであり，z_{ij}が小さいときに（すなわち無視しても全体の値に与える影響は小さいものに対して）良い性能を示すことから，実験した全てのインスタンスについて，近似解，期待値の双方に対して最も性能の良

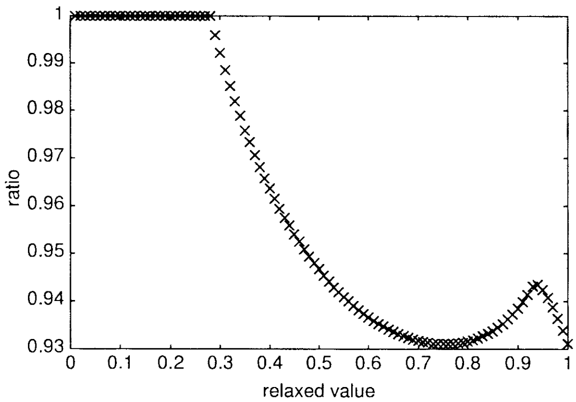

図３ z_{ij} に対する Feige-Goemans の関数の性能

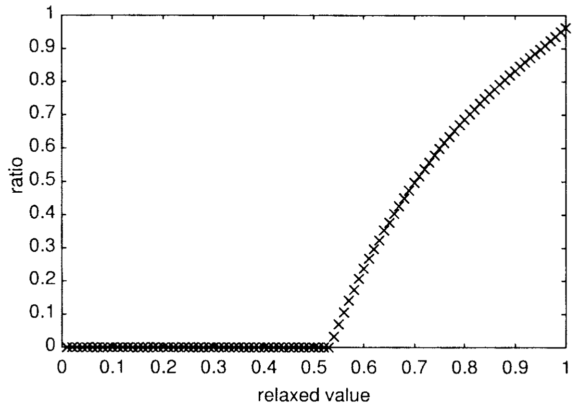

図４ z_{ij} に対する Zwick の関数の性能

かったアルゴリズムは，Zwick のアルゴリズムであった。Zwick によるアルゴリズムは，Feige-Goemans のアルゴリズムに比べて，ϵ の値が小さいインスタンスに優れているという理論的性能を持っている。しかし，実験の結果では，一般のインスタンスに対しても良い性能を示すことが観察できた。

SDP 緩和に基づく近似アルゴリズムの性能改善には，図５のような最適解の z_{ij} の分布を観察して，乱数使用ラウンディングの期待値ができるだけ大きくなるように最適緩和解のベクトルの回転を定めるというヒューリスティックスが有効であろう。

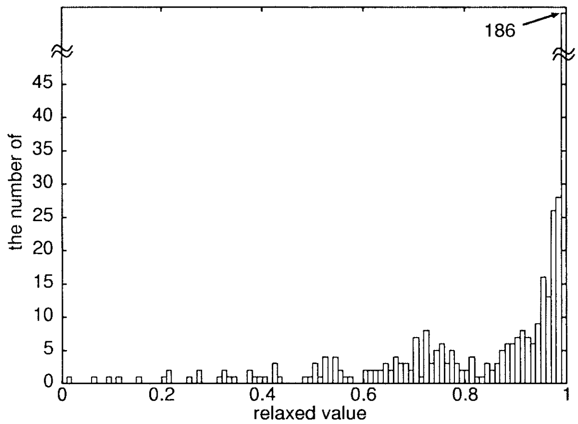

図５ 450 個の論理和の緩和値 z_{ij} の頻度

参 考 文 献

［１］ Asano T. : "Approximation algorithms for MAX SAT : Yannakakis vs. Goemans-Williamson", *Proc. 5th Israel Symposium on Theory of Computing and Systems*, pp. 24-37, 1997.

［２］ Asano T. and Williamson D. P. : "Improved approximation algorithms for MAX SAT", *Proc. 11th ACM-SIAM Symposium on Discrete Algorithms*, pp. 96-105, 2000.

［３］ Feige U. and Goemans M. X. : "Approximating the value of two prover proof systems, with applications to MAX 2SAT and MAX DICUT", *Proc. 3rd Israel Symposium on Theory of Computing and Systems*, pp. 182-189, 1995.

［４］ Goemans M. X. and Williamson D. P. : "Improved approximation algorithms for the maximum cut and satisfiability problems using semidefinite programming", *Journal of ACM* 42, pp. 1115-1145, 1995.

［５］ Halperin E. and Zwick U. : "Approximation algorithms for MAX 4SAT and rounding procedures for semidefinite programs", *Proc. 7th Conference on Integer Programming and Combinatorial Optimization, LNCS 1610*, Springer, pp. 202-217, 1999.

［６］ 岩間健一郎，浅野孝夫 : "MAX 2SAT に対する近似アルゴリズムの実際的性能評価"，情報処理学会研究報告，アルゴリズム AL-72-7, pp. 49-54, 2000.

［７］ Karloff H. and Zwick U. : "A 7/8-approximation algorithm for MAX 3SAT?", *Proc. 38th IEEE Symposium on Foundations of Computer Science*, pp. 406-415, 1997.

［８］ Zwick U. : "Finding almost-satisfying assignments", *Proc. 29th ACM Symposium on Theory of Computing*, pp. 551-560, 1998.

［９］ Zwick U. : "Outward rotations : a tool for rounding solutions of semidefinite programming relaxations, with applications to MAX CUT and other problems", *Proc. 31st ACM Symposium on Theory of Computing*, pp. 679-687, 1999.