Regular Paper

An Approximation Algorithm for MAX 3SAT

TAKAO ONO,† TOMIO HIRATA† and TAKAO ASANO††

In this paper we present a 0.769-approximation algorithm for MAX 3SAT. It is a restricted version of MAX SAT, in which every clause has at most three literals. The best previously know approximation algorithm for MAX SAT had 0.755-approximation ratio, which was given by Goemans and Williamson. Thus, we make a slight improvement by limiting MAX SAT to MAX 3SAT. Since approximating MAX 3SAT within 38/39 is NP-complete, our result means that the best approximation ratio is between 0.769 and 38/39.

gorithm.

1. Introduction

We consider the Maximum 3-Satisfiability Problem (MAX 3SAT), a restricted version of the Maximum Satisfiability Problem (MAX SAT). For MAX SAT, 0.75-approximation algorithms and 0.755-approximation algorithm were proposed. Yannakakis¹⁰⁾ first gave a 3/4-approximation algorithm for MAX SAT. Goemans and Williamson⁵⁾ gave another 3/4approximation algorithm, and reported a slight improvement of the ratio to 0.755 later4). But no good approximation algorithms for MAX 3SAT are known. In this paper we present a randomized 0.769-approximation algorithm for MAX 3SAT. Thus we make a slight improvement by limiting MAX SAT to MAX 3SAT.

Since MAX 3SAT is MAX SNP-complete⁸, for some constant $\alpha < 1$ no α -approximation algorithms for MAX 3SAT exist unless P = NP. In fact, approximating MAX 3SAT within the approximation ratio $38/39^{\pm\pm}$ is known to be NP-complete³). Our result means that the best approximation ratio is between 0.769 and 38/39.

Our result is inspired by the recent work of Goemans and Williamson⁴⁾ who used an algorithm for semidefinite optimization problems to obtain an approximation algorithm for MAX 2SAT. Their algorithm is 0.878-approximation. But in the case of MAX 3SAT, we have to treat polynomials of degree 4 instead of those of degree 2. In our algorithm we carefully truncate some portion of the polynomials. Furthermore, we combine this algorithm with other known algorithms to obtain a good approximation al-

In Section 2, some definitions are given. We present an approximation algorithm for MAX 3SAT based on the semidefinite programming in Section 3. In Section 4, we analyze the approximation ratio of the algorithm. Combining this algorithm with Johnson's algorithm⁶) and LP-relaxation algorithm⁵), we show that the 0.769-approximation ratio can be achieved in Section 5. Conclusion with some remarks is given in Section 6.

2. Definitions

Let $\{x_1, \ldots, x_n\}$ be a set of variables. A literal is either a variable x_i or its negation \bar{x}_i . A clause is a disjunction of literals. An instance I of MAX 3SAT is a collection of clauses C_1, \ldots, C_L with the corresponding positive weights w_1, \ldots, w_L , where each clause has at most three literals. We assume that no variables appear more than once in a clause, that is, we do not allow a clause like $\{x_1, \bar{x}_1, x_2\}$. We denote by $|C_j|$ the number of literals in C_j . I_j^+ and I_j^- denote the sets of indices of the variables that appear positively and negatively in C_j , respectively. The weight of an assignment is the sum of the weights of the clauses which are satisfied by the assignment.

For an algorithm A, $W_A(I)$ denotes the weight of the assignment produced by A for an input I. The approximation ratio of A is α if $W_A(I)$ is at least α times the weight $W_{OPT}(I)$ of the optimal assignment for any instance I. We call A an α -approximation algorithm. We use the notation f(x) for multivariate func-

[†] School of Engineering, Nagoya University

^{††} Department of Information and System Engineering, Chuo University

[★] This is a revised version of Ref. 7), in which we erroneously claimed that our algorithm achieves 0.80-approximation.

^{★★} At present, the ratio is reduced to 26/27²).

tion $f(x_1,\ldots,x_n)$.

3. An Approximation Algorithm

In this section, we present an approximation algorithm for MAX 3SAT. This algorithm solves MAX 3SAT in the same way as Goemans and Williamson's algorithm for MAX 2SAT. The algorithm consists of the following steps.

- (1) Translate an instance of MAX 3SAT into an integer programming problem.
- (2) Reduce the problem to a semidefinite programming problem with relaxation and solve it.
- (3) Construct a truth assignment of the original MAX 3SAT problem from a solution to the semidefinite programming problem.

We explain each of the above steps in the following subsections.

3.1 Translation of an Instance into an Integer Programming Problem

We arithmetize the clause C_j as

$$C_j(x) = 1 - \prod_{i \in I_j^+} (1 - x_i) \prod_{i \in I_j^-} x_i.$$

Thus, we can formulate MAX 3SAT by the following integer programming problem:

$$\begin{array}{ll} \text{Maximize} & \sum_{j=1}^L w_j C_j(\boldsymbol{x}) \\ \text{subject to} & x_i \in \{0,1\} & (1 \leq i \leq n). \end{array}$$

Because $C_j(x) \leq \sum_{i \in I_j^+} x_i + \sum_{i \in I_j^-} (1 - x_i)$, we use the following formulation of MAX 3SAT.

$$\begin{aligned} \text{Maximize} \quad & \sum_{j=1}^L w_j z_j \\ \text{subject to} \\ & z_j \leq C_j(\boldsymbol{x}) \\ & z_j \leq \sum_{i \in I_j^+} x_i + \sum_{i \in I_j^-} (1 \leq j \leq L), \\ & 0 \leq z_j \leq 1 \\ & x_i \in \{0,1\} \end{aligned}$$

In the rest of this paper we omit the ranges of i and j if they are clear from the context.

Next we introduce variables y_i (i = 0, 1, ..., n), whose values are +1 or -1 and replace x_i with $(1 + y_0y_i)/2$. Thus MAX 3SAT can be written as:

Maximize
$$\sum_{j=1}^{L} w_j z_j$$

subject to
$$z_j \leq C_j'(\boldsymbol{y}),$$
 $z_j \leq \sum_{i \in I_j^+} \frac{1 + y_0 y_i}{2}$ $+ \sum_{i \in I_j^-} \frac{1 - y_0 y_i}{2},$ $0 \leq z_j \leq 1,$ $y_i \in \{-1, +1\},$

where $C'_j(y) = C_j((1 + y_0y_1)/2,...,(1 + y_0y_n)/2).$

Now we consider the case that C_j has three literals, say, x_i , x_j and x_k . In this case,

$$C'_{j}(\mathbf{y}) = \frac{1}{8} [(1 + y_{0}y_{i}) + (1 + y_{0}y_{j}) + (1 + y_{0}y_{k}) + (1 - y_{i}y_{j}) + (1 - y_{j}y_{k}) + (1 - y_{k}y_{i}) + (1 + y_{0}y_{i}y_{j}y_{k})].$$

Here, if there is a negative literal in those three literals, we replace the corresponding variable y with -y. We denote by $c_j(y)$ and $d_j(y)$ the sum of the first six terms and the last term in $C'_j(y)$, respectively, that is, $C'_j(y) = c_j(y) + d_j(y)$ and $c_j(y)$ contains terms of degree at most 2. The next lemma bounds the value of $C'_j(y)$.

Lemma 1 For all C_j containing three literals, $c_j(y) \leq C'_j(y) \leq (4/3)c_j(y)$.

Proof The first inequality is trivial because $d_j(\boldsymbol{y}) \geq 0$. The second inequality comes as follows. Consider that clause C_j has three literals. If $d_j(\boldsymbol{y}) = 0$, then $C'_j(\boldsymbol{y}) = c_j(\boldsymbol{y}) \geq 0$ and thus $(4/3)c_j(\boldsymbol{y}) \geq C'_j(\boldsymbol{y})$. We assume that $d_j(\boldsymbol{y}) \neq 0$. This means that $d_j(\boldsymbol{y}) = 1/4$ and $C'_j(\boldsymbol{y}) \neq 0$. Because $C'_j(\boldsymbol{y}) \in \{0,1\}$, $C'_j(\boldsymbol{y})$ must equal 1. Thus, $c_j(\boldsymbol{y}) = 3/4$ and $C'_j(\boldsymbol{y}) \leq (4/3)c_j(\boldsymbol{y})$.

 $C'_j(\boldsymbol{y}) \leq (4/3)c_j(\boldsymbol{y}).$ We define $c_j(\boldsymbol{y}) = C'_j(\boldsymbol{y})$ for all the clauses C_j which have less than three literals. Thus we can reformulate MAX 3SAT as follows:

Maximize
$$\sum_{j=1}^{L} w_{j} z_{j}$$

subject to $z_{j} \leq c_{j}(\mathbf{y})$ if $|C_{j}| \leq 2$, $z_{j} \leq \frac{4}{3} c_{j}(\mathbf{y})$ if $|C_{j}| = 3$, $z_{j} \leq \sum_{i \in I_{j}^{+}} \frac{1 + y_{0} y_{i}}{2} + \sum_{i \in I_{j}^{-}} \frac{1 - y_{0} y_{i}}{2}$, $0 \leq z_{j} \leq 1$, $y_{i} \in \{-1, +1\}$. (1)

3.2 Reducing to a Semidefinite Programming Problem

In the problem (1), $c_j(y)$ contains the constant terms and the product terms of two variables. This fact enables us to relax the problem by using the following replacements:

- (1) replace variables y_i with vectors v_i whose norms are 1,
- (2) replace the product $y_i y_j$ with the inner product of the corresponding vectors $v_i \cdot v_j$,
- (3) replace the inner product $v_i \cdot v_j$ with a new variable y_{ij} .

Let V be the matrix (v_0, v_1, \ldots, v_n) and let V^T be its transpose. Then the $(n+1) \times (n+1)$ matrix $Y = (y_{ij})$ equals V^TV , which is symmetric and positive semidefinite. Thus we denote by $c_j(Y)$ a relaxed version of $c_j(y)$ and we have the maximization problem:

Maximize
$$\sum_{j=1}^{L} w_{j} z_{j}$$
subject to
$$z_{j} \leq c_{j}(Y) \quad \text{if } |C_{j}| \leq 2,$$

$$z_{j} \leq \frac{4}{3} c_{j}(Y) \quad \text{if } |C_{j}| = 3,$$

$$z_{j} \leq \sum_{i \in I_{j}^{+}} \frac{1 + y_{0i}}{2}$$

$$+ \sum_{i \in I_{j}^{-}} \frac{1 - y_{0i}}{2},$$

$$0 \leq z_{j} \leq 1,$$

$$y_{ii} = 1,$$
matrix Y is symmetric and positive semidefinite.

An optimization problem which is represented by a linear system of the entries of a symmetric and positive semidefinite matrix is called a "semidefinite programming problem." It is well-known that a semidefinite programming problem can be solved in polynomial time with respect to the input size and $\log(1/\epsilon)$ within any additive error ϵ^{1} . See Goemans and Williamson⁴) for details and references.

For (2) we consider the matrix

where diag
$$(z_1, z_2, ..., z_L)$$
 whose diagonal entries are $z_1, z_2, ..., z_L$

where $\operatorname{diag}(z_1, z_2, \ldots, z_L)$ is a diagonal matrix whose diagonal entries are z_1, z_2, \ldots, z_L . Then the objective function and all constraints in (2) are linear combinations of the entries of \tilde{Y} , and \tilde{Y} is symmetric and positive semidefinite. Thus,

(2) is a semidefinite programming, and hence we can find its solution (Y^*, z^*) (with a negligibly small error) in polynomial time.

3.3 Finding Approximate Solutions to MAX 3SAT

We use the following scheme, which is used in Goemans and Williamson⁴⁾, to obtain an approximate assignment x^S of MAX 3SAT:

First we obtain the vectors v_i^* by Cholesky decomposition of Y^* . Then we take the random vector \boldsymbol{r} whose norm is 1 and let $y_i^S = \text{sign}(\boldsymbol{r} \cdot \boldsymbol{v}_i^*)$. x_i^S is set to $(1 + y_0^S y_i^S)/2$.

4. Analysis of the Approximation Ratio

In this section, we show the weight of the approximate assignment x^S .

Lemma 2 The weight of x^S is bounded as follows.

$$\begin{split} \sum_{j=1}^L w_j C_j(\boldsymbol{x}^S) &\geq \alpha \sum_{|C_j| \leq 2} w_j z_j^* \\ &+ \frac{3\alpha}{4} \sum_{|C_j| = 3} w_j z_j^*, \end{split}$$

where $\alpha = \min_{0 \le \theta \le \pi} (\theta/\pi)/[(1 - \cos \theta)/2] = 0.87856...$

To prove this lemma, we use the next lemma⁴⁾. Proof is given so that this paper may be self-contained.

Lemma 3 For α in Lemma 2,

$$E[(1 - y_i y_j)/2] \ge \alpha (1 - y_{ij}^*)/2,$$

$$E[(1 + y_i y_j)/2] \ge \alpha (1 + y_{ij}^*)/2.$$

Proof We can prove both inequalities in the same manner, so here we prove only the first one. Because the value of $(1 - y_i y_j)/2$ is either 0 or 1,

$$E[(1 - y_i y_j)/2]$$

$$= \Pr\{y_i \neq y_j\}$$

$$= \Pr\{\operatorname{sign}(\boldsymbol{r} \cdot \boldsymbol{v}_i^*) \neq \operatorname{sign}(\boldsymbol{r} \cdot \boldsymbol{v}_j^*)\}$$

$$= 2\Pr\{\boldsymbol{r} \cdot \boldsymbol{v}_i^* > 0 \text{ and } \boldsymbol{r} \cdot \boldsymbol{v}_j^* < 0\}$$
(by symmetry).

We denote by θ_{ij} the angle between v_i^* and v_j^* . The set $\{r: r \cdot v_i^* > 0 \text{ and } r \cdot v_j^* < 0\}$ is a spherical lune with angle θ_{ij} , whose measure equals $\theta_{ij}/(2\pi)$ times the measure of the full sphere. Thus

$$E[(1-y_iy_j)/2] = 2\Pr\{oldsymbol{r}\cdotoldsymbol{v}_i^*>0 ext{ and } oldsymbol{r}\cdotoldsymbol{v}_j^*<0\} = 2 heta_{ij}/(2\pi) = heta_{ij}/\pi.$$

From the definition of α , for any θ $(0 \le \theta \le \pi)$

$$heta/\pi \geq \alpha(1-\cos\theta)/2$$
. Thus
$$E[(1-y_iy_j)/2] = \theta_{ij}/\pi$$

$$\geq \alpha(1-\cos\theta_{ij})/2$$

$$= \alpha(1-y_{ij}^*)/2.$$

Proof (of Lemma 2) From Lemma 3, $E[c_j(\boldsymbol{y})] \geq \alpha c_j(Y^*)$. Thus for clauses C_j which has at most two literals

 $E[C_j(\boldsymbol{x})] = E[c_j(\boldsymbol{y})] \geq \alpha c_j(Y^*) \geq \alpha z_j^*.$ On the other hand, for the clauses C_j which has

 $E[C_j(\boldsymbol{x})] \geq E[c_j(\boldsymbol{y})] \geq \alpha c_j(Y^*) \geq 3\alpha z_j^*/4.$ Combining these inequalities,

$$\begin{split} &\sum_{j=1}^{L} w_{j} C_{j}(\boldsymbol{x}^{S}) \\ &\geq E \left[\sum_{j=1}^{L} w_{j} C_{j}(\boldsymbol{x}) \right] \\ &= \sum_{j=1}^{L} w_{j} E[C_{j}(\boldsymbol{x})] \\ &= \sum_{|C_{j}| \leq 2} w_{j} E[C_{j}(\boldsymbol{x})] + \sum_{|C_{j}| = 3} w_{j} E[C_{j}(\boldsymbol{x})] \\ &\geq \sum_{|C_{j}| \leq 2} w_{j} \alpha z_{j}^{*} + \sum_{|C_{j}| = 3} w_{j} \cdot 3\alpha z_{j}^{*} / 4 \\ &= \alpha \sum_{|C_{j}| \leq 2} w_{j} z_{j}^{*} + \frac{3\alpha}{4} \sum_{|C_{j}| = 3} w_{j} z_{j}^{*}. \end{split}$$

5. 0.769-Approximation Algorithm for MAX 3SAT

In this section, we consider the approximation algorithm A which selects one of three assignments x^J , x^L and x^S with some probabilities. x^S is obtained by using the rounding scheme described in Section 3. x^J and x^L are obtained by the random assignment method9), that is, each variable x_i is independently set to 1 with probability p_i .

For x^J , we let $p_i = 1/2$ for all i. For x^L , we let $p_i = (1 + y_{0i}^*)/2$. Thus for the weights of x^J and x^L , we can show the following lemmas.

Lemma 4 The weight of x^J is bounded as

$$\sum_{j=1}^{L} w_j C_j(\boldsymbol{x}^J) \ge \sum_{k=1}^{3} \left[\left(1 - \frac{1}{2^k} \right) \sum_{|C_j| = k} w_j \right].$$

Proof Because we consider that $p_i = 1/2$ for all i, the probability that the clause C_j is satisfied is $1 - 1/2^{|C_j|}$. Thus,

$$\sum_{j=1}^{L} w_j C_j(\boldsymbol{x}^J) \ge E \left[\sum_{j=1}^{L} w_j C_j(\boldsymbol{x}) \right]$$

$$= \sum_{j=1}^{L} w_j E[C_j(\boldsymbol{x})]$$

$$= \sum_{j=1}^{L} w_j \left(1 - \frac{1}{2^{|C_j|}} \right)$$

$$= \sum_{k=1}^{3} \left[\left(1 - \frac{1}{2^k} \right) \sum_{|C_j| = k} w_j \right].$$

Lemma 5 The weight of x^L is bounded as follows.

$$\sum_{j=1}^{L} w_j C_j(x^L) \ge \sum_{k=1}^{3} \left[eta_k \sum_{|C_j|=k} w_j z_j^*
ight],$$

where $\beta_k = 1 - (1 - 1/k)^k$.

Proof We can see that (2) is a version of LP relaxation of MAX SAT⁵). Thus, we can apply lemma 3.1 in Ref. 5) to \boldsymbol{x}^L to obtain the bound of its weight:

$$\sum_{j=1}^{L} w_j C_j(\boldsymbol{x}^L) \ge \sum_{j=1}^{L} \beta_{|C_j|} w_j z_j^*,$$
where $\beta_k = 1 - (1 - 1/k)^k$. Restricting MAX

SAT to MAX 3SAT, we can obtain the lemma.□

Finally we show the approximation ratio of algorithm A. We assume that A selects x^{J} , x^L and x^S with the probabilities p_J , p_L , p_S , respectively.

Theorem 1 Algorithm A is 0.769-approximation for appropriate p_J , p_L and p_S .

Proof Algorithm A is a-approximation for the value of a satisfying

$$\begin{array}{llll} \frac{1}{2} \; p_J + & p_L + \; \alpha \; p_S \; \geq \; a, \\ \\ \frac{3}{4} \; p_J + \; \frac{3}{4} \; p_L + \; \alpha \; p_S \; \geq \; a, \\ \\ \frac{7}{8} \; p_J + \frac{19}{27} \; p_L + \frac{3\alpha}{4} \; p_S \; \geq \; a. \end{array}$$

These inequalities express the expectations of $C_j(x)$ corresponding to the clauses with one, two and three literals, respectively. The upper bound of a, obtained by solving a linear programming problem, is 0.7694.... Thus algorithm A is 0.769-approximation.

6. Concluding Remarks

We have presented a randomized 0.769-

approximation algorithm for MAX 3SAT. The best previously known approximation algorithm for MAX 3SAT had 0.755 approximation ratio as a special case of the same bound for MAX SAT, so we have made a slight improvement by limiting MAX SAT to MAX 3SAT.

For MAX 3SAT, approximating within the ratio 38/39 is known to be NP-complete. The approximation ratio of our algorithm is much smaller than 38/39. A challenging problem is to narrow this gap.

Acknowledgments The authors would like to thank Michel Goemans for his comments on the first version of this paper, and anonymous referees for their valuable comments which improve the exposition of this paper.

References

- 1) Alizadeh, F.: Interior Point Methods in Semidefinite Programming with Applications to Combinatorial Optimization, SIAM J. Optimization, Vol.5, No.1, pp.13-51 (1995).
- Bellare, M., Goldreich, O. and Sudan, M.: Free Bits, PCPs and Non-Approximability— Towards Tight Results, Unpublished Manuscript.
- 3) Bellare, M., Goldreich, O. and Sudan, M.: Free Bits, PCPs and Non-Approximability—Towards Tight Results, TR95-024, ECCC, the Electronic Colloquium on Computational Complexity (1995).
- 4) Goemans, M.X. and Williamson, D.P.: .878-Approximation Algorithms for MAX CUT and MAX 2SAT, *Proc. 26th STOC*, pp.422-431 (1994).
- 5) Goemans, M.X. and Williamson, D.P.: New 3/4-Approximation Algorithms for the Maximum Satisfiability Problem, SIAM J. Disc. Math., Vol.7, No.4, pp.656-666 (1994).
- Johnson, D.S.: Approximation Algorithms for Combinatorial Problems, J. Comput. Syst. Sci., Vol.9, pp.256-278 (1974).
- 7) Ono, T., Hirata, T. and Asano, T.: An Approximation Algorithm for MAX 3SAT, *Proc.* 6th ISAAC, pp.163-170 (1995).
- Papadimitriou, C.H. and Yannakakis, M.: Optimization, Approximation, and Complexity Classes, *Proc. 20th STOC*, pp.229-234 (1988).
- 9) Raghavan, P. and Thompson, C.D.: Randomized Rounding: A Technique for Provably Good

- Algorithms and Algorithmic Proofs, Combinatorica, Vol.7, No.4, pp.365-374 (1987).
- 10) Yannakakis, M.: On the Approximation of Maximum Satisfiability, *Journal of Algorithms*, Vol.17, No.3, pp.475-502 (1994).

(Received May 22, 1966) (Accepted July 4, 1996)



Takao Ono was born in 1970. He received the M.E. degree from Nagoya University in 1995. He is currently studying towards the Ph.D. degree in the Graduate School of Engineering in the same university. His research in-

terests include approximation algorithms.



Tomio Hirata was born in 1949. He received the B.S., M.S. and Ph.D. in Computer Science, all from Tohoku University in 1976, 1978, and 1981, respectively. He is currently a Professor in Electronics Depart-

ment, Nagoya University. From 1981 to 1986 he was an Assistant Professor of the Computer Science Department, Toyohashi University of Technology. His research interests include graph algorithms, data structures, and parallel algorithms.



Takao Asano was born in Miyagi Prefecture on January 7, 1949. He received the B.E., M.E. and D.E. degrees in Electrical Engineering and Communications from Tohoku University, Sendai, Japan, in 1972,

1974 and 1977, respectively. From 1977 to 1980, he was Research Associate of Tohoku University. From 1980 to 1985, he was Lecturer of Department of Mathematical Engineering of the University of Tokyo. From 1985 to 1992, he was Associate Professor of Sophia University. From 1992, he is Professor of Department of Information and System Engineering of Chuo University. His research interests include network algorithms and computational geometry.