CMPUT 498/501: Advanced Algorithms **Fall 2025**

Lecture 18 (Nov 5, 2025): Approximation Algorithms for Max SAT

Lecturer: Mohammad R. Salavatipour *Scribe: Kaining Yang*

18.1 Introduction to Maximum Satisfiability

In this lecture, we study approximation algorithms for the Maximum Satisfiability (Max SAT) problem, which is a natural optimization variant of the classic SAT decision problem. Unlike SAT, where we simply ask whether there exists an assignment satisfying all clauses, Max SAT try to maximize the total weight of satisfied clauses even when satisfying all clauses simultaneously may be impossible.

18.1.1 Problem Formulation

We are given a Boolean formula $\mathcal{F}$ in conjunctive normal form (CNF) over n variables $x_1, \ldots, x_n$ with m clauses $C_1, \ldots, C_m$. Each clause C_j has an associated weight $w_j > 0$. Our goal is to find a truth assignment to the variables that maximizes the total weight of satisfied clauses, denoted by $\sum_{j:C_j \text{ is satisfied}} w_j$.

When all weights are uniform (i.e., $w_j = 1$ for all j), the objective simply becomes maximizing the number of satisfied clauses. This special case is often of particular interest in practice.

We also consider restricted versions of the problem:

- **Max-k-SAT:** Every clause contains at most k literals.
- **Max-E-k-SAT:** Every clause contains exactly k literals.

Despite the fact that 2SAT is solvable in polynomial time, the optimization version becomes hard immediately:

Theorem 1 *Max-k-SAT is NP-hard for any $k \geq 2$.*

This hardness result motivates our study of approximation algorithms for Max-SAT.

18.2 Simple Randomized Algorithm for Max-SAT

This is perhaps the most obvious randomized algorithm: for each variable x_i, independently flip a fair coin and set it to True or False with equal probability $\frac{1}{2}$. Despite its simplicity, this algorithm already achieves a non-trivial approximation guarantee.

Theorem 2 (Johnson'74) *The randomized algorithm gives a $\frac{1}{2}$-approximation for Max-SAT.*

Proof. Let $\mathcal{C}$ be the random truth assignment and define

$$Y_j = \begin{cases} 1 & \text{if } C_j \text{ is sat} \\ 0 & \text{otherwise} \end{cases}$$

Let $W = \sum_{j=1}^{m} w_j Y_j$ (total weight of satisfied clauses).

$$\begin{aligned} \mathbb{E}[W] &= \sum_{j=1}^{m} w_j \mathbb{E}[Y_j] = \sum_{j=1}^{m} w_j \cdot \Pr[C_j \text{ is sat}] \\ &= \sum_j w_j \left(1 - \left(\frac{1}{2}\right)^{|C_j|}\right) \\ &\geq \frac{1}{2} \sum_j w_j \\ &\geq \frac{1}{2} \text{OPT} \end{aligned}$$

Note that if $|C_j| \geq k$ for all j, then this is a $\left(1 - \frac{1}{2^k}\right)$-approximation, which gives a $\frac{7}{8}$-approximation for Max-E3SAT. ■

Theorem 3 (Håstad'99) *There is no $\left(\frac{7}{8} - \varepsilon\right)$-approximation for any fixed $\varepsilon > 0$ for Max-E3SAT unless $P = NP$.*

18.3 Derandomization Using Method of Conditional Expectation

Lemma 1 *Suppose we have assigned values $a_1, \ldots, a_i$ to variables $x_1, \ldots, x_i$. Then we can compute the expected value of the solution if the rest of the variables are assigned randomly.*

Proof. Let $\mathcal{F}'$ be the formula on variables $x_{i+1}, \ldots, x_n$ obtained from $\mathcal{F}$ by substituting the values of $x_1, \ldots, x_i$ and simplifying $\mathcal{F}$. Clearly the expected value of solution for any truth assignment to $\mathcal{F}'$ can be computed in polynomial time.

Add this to the weight of clauses that are already satisfied by $x_1, \ldots, x_i$ and we obtain the expected value of $\mathcal{F}$ given $x_1, \ldots, x_i$ are fixed. ■

This suggests a simple deterministic algorithm: Consider x_1. For each of T/F assignment to x_1, compute the expected value $\mathbb{E}[W|x_1 = T]$ and $\mathbb{E}[W|x_1 = F]$, whichever is bigger, assign x_1 accordingly. Do this for all variables iteratively. Say this value is v for x_1.

$$\mathbb{E}[W] = \mathbb{E}[W|x_1 = T] \cdot \Pr[x_1 = T] + \mathbb{E}[W|x_1 = F] \cdot \Pr[x_1 = F] = \frac{1}{2}(\mathbb{E}[W|x_1 = T] + \mathbb{E}[W|x_1 = F])$$

If we set x_1 as in above, then $\mathbb{E}[W|x_1 = v] \geq \mathbb{E}[W] \geq \frac{1}{2}\text{OPT}$.

18.4 Randomized Algorithm with Biased Coin

First assume we have an instance where all clauses of length 1 have non-negative literals (i.e. no clause like $\overline{x}_j = C_j$). We set each variable $x_i = T$ with probability P where $P > \frac{1}{2}$ (to be determined).

For any C_j:

- If $|C_j| = 1$, then $\Pr[C_j \text{ is sat}] = P$
- If $|C_j| \geq 2$, then $\Pr[C_j \text{ is sat}] = 1 - P^\alpha(1-P)^\beta \geq 1 - P^{\alpha+\beta} \geq 1 - P^2$, where α is the number of negative literals and β is the number of non-negative literals.

Lemma 2 $\Pr[C_j \textit{ is sat}] \geq \min\{P, 1-P^2\}$

Set $P = 1 - P^2 \Rightarrow P = \frac{\sqrt{5}+1}{2} \approx 0.618$

$$\mathbb{E}[W] = \sum_j w_j \cdot \Pr[C_j \text{ is sat}] \geq P \sum_j w_j \geq P \cdot \text{OPT}$$

Theorem 4 *If all clauses of length 1 have positive literals, then this is a P-approximation where $P = \frac{\sqrt{5}+1}{2}$.*

What if there are 1-clauses with negative literals? For example, $C_j = \overline{x}_i$. We can easily replace $\overline{x}_i$ with x_i' and x_i with $\overline{x}_i'$.

The only problem is when both x_i and $\overline{x}_i$ appear as a 1-clause, e.g., $C_j = x_i$ and $C_\ell = \overline{x}_i$.

Without loss of generality, assume $w_j \geq w_\ell$. Define $w_i' = w_\ell$. Then

$$\text{OPT} \leq \sum_j w_j - \sum_{i\text{:1-clause with negtive literal \& positive 1-clause}} w_i'$$

Let U denote the set of indices of clauses, excluding those with negative literals.

We can still show:

$$\mathbb{E}[W] = \sum_j w_j \cdot \Pr[C_j \text{ is sat}] \geq \sum_{j\in U} w_j \cdot \Pr[C_j \text{ is sat}] \geq P \sum_{j \in U} w_j \geq P \cdot \text{OPT}$$

Theorem 5 *Biased coin implies a P-approximation for Max-SAT.*

18.5 LP-rounding for Max-SAT

Let $P_j(N_j)$ be the set of variables that appear in clause C_j in positive (negative) form, and $\ell_j = |P_j| + |N_j|$. For any x_i, we have $y_i \in \{0,1\}$ where $y_i = 1(0)$ iff $x_i = T(F)$.

Also for each C_j, we have $z_j = 1$ iff C_j is satisfied.

$$\max \sum_j w_j \cdot z_j \qquad \text{(Random Rounding for Max-SAT)}$$

subject to:

$$\sum_{i\in P_j} y_i + \sum_{i\in N_j}(1-y_i) \geq z_j \qquad \forall C_j \in C$$
$$0 \leq z_i \cdot y_i \leq 1$$

Random Rounding: Solve the LP, let (y^*, z^*) be the optimal solution. For each x_i, set it to T with probability y_i^*. Let $\hat{x}$ be the integer solution obtained.

18.5.1 Analysis using Goemans-Williamson Technique

Theorem 6 (Goemans/Williamson '94) *R.R. is a $(1-\frac{1}{e})$-approximation for Max-SAT (≈ 0.632).*

Proof. We use the following two facts:

Fact 1: Arithmetic/Geometric mean inequality: for non-negative numbers $a_1, \ldots, a_k$:

$$\frac{a_1 + \ldots + a_k}{k} \geq \sqrt[k]{a_1 a_2 \cdots a_k} \qquad \left(\frac{\sum a_i}{k}\right)^k \geq \prod_i a_i$$

Fact 2: If $f(x)$ is concave over $[0,1]$ (i.e., $f''(x) \leq 0$) and $f(0) = 0$ and $f(1) = \alpha$, then the function f is lower bounded by the line going through $(0,0)$ and $(1,\alpha)$.

Let W_j be the weight contributed by C_j to the total. We first prove a lemma. ■

Lemma 3 *For each C_j: $\mathbb{E}[W_j] \geq \left[1-\left(1-\frac{1}{k}\right)^k\right] w_j \cdot z_j^*$ for any clause C_j with $\ell_j = k$.*

Proof.

$$\begin{aligned}
\Pr[C_j \text{ is sat}] &= 1 - \prod_{i\in P_j}(1-y_i^*) \prod_{i\in N_j} y_i^* \\
&\geq 1 - \left(\frac{\sum_{i\in P_j}(1-y_i^*) + \sum_{i\in N_j} y_i^*}{k}\right)^k \qquad \text{(use A.G.M)} \\
&\geq 1 - \left(1 - \frac{\sum_{i\in P_j} y_i^* + \sum_{i\in N_j}(1-y_i^*)}{k}\right)^k
\end{aligned}$$

By constraint for C_j:

$$\geq 1 - \left(1 - \frac{z_j^*}{k}\right)^k$$

Consider function $g(z) = 1 - \left(1-\frac{z}{k}\right)^k$, where $g(0) = 0$ and $g(1) = 1 - \left(1-\frac{1}{k}\right)^k$, and $g(\cdot)$ is concave in $[0,1]$.

By Fact 2: $g(z) \geq \left(1 - \left(1-\frac{1}{k}\right)^k\right) \cdot z$

So $\Pr[C_j \text{ is sat}] \geq \left[1 - \left(1 - \frac{1}{k}\right)^k\right] z_j^*$

$$\mathbb{E}[W] = \sum_j w_j \cdot \Pr[C_j \text{ is sat}] \geq \left[1 - \left(1 - \frac{1}{k}\right)^k\right] \sum_j w_j \cdot z_j^* \geq \left(1 - \frac{1}{e}\right) \text{OPT}$$

Since $k \to \infty$, $\left(1 - \frac{1}{k}\right)^k \to \frac{1}{e}$.

■

Remark: This algorithm performs better than $1 - \frac{1}{e}$ if C_j's are small.

Remark: We can also derandomize this.

18.6 A $\frac{3}{4}$-approximation for Max-SAT

Take the better of:

- Randomized Algorithm (suppose we set $a = 0/1$ with probability $\frac{1}{2}$ and)
- Random Rounding algorithm (LP) (if $a = 1$)

Let $\mathcal{Z}^*$ be the optimal value of LP and W_j be the contribution of clause C_j in our algorithm.

Lemma 4 *For each C_j: $\mathbb{E}[W_j] \geq \frac{3}{4} w_j \cdot z_j^*$*

Proof. Suppose C_j has k literals, and define $\alpha_k = 1 - \frac{1}{2^k}$ and $\beta_k = 1 - \left(1 - \frac{1}{k}\right)^k$.

$$\mathbb{E}[W_j | a = 0] \geq \left(1 - \frac{1}{2^k}\right) w_j \geq \alpha_k w_j \cdot z_j^* \qquad (\text{Since } z_j^* \leq 1)$$

$$\mathbb{E}[W_j | a = 1] \geq \left[1 - \left(1 - \frac{1}{k}\right)^k\right] w_j \cdot z_j^* = \beta_k w_j \cdot z_j^*$$

$$\begin{aligned} \mathbb{E}[W_j] &= \mathbb{E}[W_j | a = 0] \cdot \Pr[a = 0] + \mathbb{E}[W_j | a = 1] \cdot \Pr[a = 1] \\ &\geq \frac{1}{2}[\alpha_k + \beta_k] w_j \cdot z_j^* \end{aligned}$$

$\alpha_1 + \beta_1 = \alpha_2 + \beta_2 = \frac{3}{2}$, for $k \geq 3$, $\alpha_k + \beta_k \geq \frac{7}{8} + (1 - \frac{1}{e}) \geq \frac{3}{2}$

So $\mathbb{E}[W] = \sum_j \mathbb{E}[W_j] \geq \frac{3}{4} \sum_j w_j \cdot z_j^* \geq \frac{3}{4}\text{OPT}$ ■

We can also derandomize this algorithm.

18.6.1 Tight Example for LP (Integrality Gap)

$(x_1 \vee x_2) \wedge (\overline{x}_1 \vee x_2) \wedge (x_1 \vee \overline{x}_2) \wedge (\overline{x}_1 \vee \overline{x}_2)$ where all $w_j = 1$.

$\mathrm{OPT} = 3$, but if we set $y_i = \frac{1}{2}$ and $z_j = 1$, we have a feasible LP with value 4.

This gives an integrality gap of $\frac{4}{3}$.

Best known ratio for Max-SAT: ≈ 0.7846 (might be 0.8331-approximation based on a plausible conjecture).