Max SAT

collection of (AND, $\wedge$) "clauses"
literals (x_i, $\bar{x}_i$) connected by OR ($\vee$)

Input: CNF formula ϕ (clauses can have any # literals)
on variables $X = \{x_1 \dots x_n\}$.
(i.e., $\phi = (x_1) \wedge (\bar{x}_1 \vee x_2) \wedge (\bar{x}_3 \vee x_4 \vee x_5)$)
(k-CNF: each clause has $\leq k$ literals)

Output: assignment $f: X \to \{T, F\}$ to maximize (# satisfied clauses)

Theorem. Max SAT is NP-hard.

Theorem. If every clause has exactly 3 literals from 3 different variables, $\exists$ $7/8$-(randomized) approximation algorithm.

Pf. Consider "naive" random $f: X \to \{T, F\}$ that assigns

$$x_i = \begin{cases} T & \text{w.p. } 1/2 \\ F & \text{w.p. } 1/2 \end{cases} \quad \text{independently.}$$

For each clause C, $\Pr[C \text{ is sat. by } f] = 7/8$.

$\therefore \mathbb{E}[(\text{\# of clauses sat. by } f)] = 7/8 \cdot (\text{\# clauses in } \phi)$ $\square$

Generalizing this, if each clause has k literals coming from k different variables, $\exists$ (randomized) $(1 - 1/2^k)$-approx. algo.

Gets better as k gets ℓ!

What happens different clauses have different sizes?

(WLOG, literals in one clause are from different variables – why?)

Clauses with one literal are the worst ones, satisfied w.p. $1/2$.

LP-Based Algorithm

Suppose $\phi = C_1 \wedge \dots \wedge C_m$. For $j \in [m]$, $S_j^+ := \{i \in [n] : x_i \in C_j\}$, $S_j^- := \{i : \overline{x_i} \in C_j\}$.

(So, $C_j = \left(\bigvee_{i \in S_j^+} x_i\right) \vee \left(\bigvee_{i \in S_j^-} \overline{x_i}\right)$.

LP relaxation (vars : $(y_i)_{i \in [n]} : y_i = 1$ indicates $x_i \leftarrow T$
$(z_j)_{j \in [m]} : z_j = 1$ " C_j is satisfied)

$$\max \sum_{j \in [m]} z_j.$$

s.t.

$$\sum_{i \in S_j^+} y_i + \sum_{i \in S_j^-} (1 - y_i) \geq z_j \quad \forall j \in [m]$$

$$0 \leq z, y \leq 1.$$

if $z_j = 1$, then at least one $i \in S_j^+$ has $y_i = 1$ or at least $i \in S_j^-$ has $y_i = 0$!

Again, randomized rounding (w.r.t. LP): $x_i \leftarrow T$ w.p. y_i independently.

Fix some $C_j \in \phi$ and let $k_j = (\#$ literals of $C_j)$.

$$\Pr[C_j \text{ unsatisfied}] = \prod_{i \in S_j^+} (1 - y_i) \prod_{i \in S_j^-} y_i.$$

(let $t = \sum_{i \in S_j^+} (1 - y_i) + \sum_{i \in S_j^-} y_i \leq k_j - z_j$)

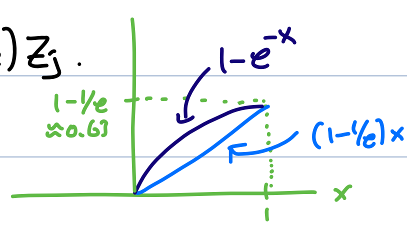
$$\leq \left(t / k_j\right)^{k_j}$$

(AM-GM: if $a_1 \dots a_k$ are positive numbers with fixed $a_1 + \dots + a_k = t$, $a_1 a_2 \cdots a_k$ is maximized when $a_i = t/k$ $\forall i$!)

$$\leq \left(1 - z_j / k_j\right)^{k_j}$$

$$\leq e^{-z_j}.$$

$$\Pr[C_j \text{ satisfied}] \geq 1 - e^{-z_j} \geq (1 - 1/e) z_j.$$



Final algorithm: Do "naive random" w.p. $1/2$ and "LP-based-random" w.p. $1/2$.

Fix some $C_j \in \phi$. $\Pr[C_j \text{ satisfied}]$?

	naive	LP rounding.	overall
$k_j = 1$	$1/2$	$\geq 1-(1-z_j)$	$\geq 1/4 + z_j/2 \geq 3/4 \cdot z_j$
$k_j = 2$	$3/4$	$\geq 1-(1-z_j/2)^2$	$\geq 3/8 + z_j/2 - z_j^2/8 \geq 3/4 \cdot z_j.$
$k_j \geq 3$	$\geq 7/8$	$\geq (1-1/e) z_j$	$\geq \underbrace{7/16}_{\substack{\shortparallel \\ 0.4375}} + \underbrace{\frac{(1-1/e)}{2}}_{\substack{\simeq \\ 0.32}} z_j \geq 3/4 \cdot z_j$

So, $\Pr[C_j \text{ satisfied}] \geq 3/4\, z_j$ for all j and

$\mathbb{E}[\# \text{ of satisfied clauses}] \geq 3/4 \cdot OPT_{LP}$!

Thm. $\forall$ constant $\varepsilon > 0$, $\nexists$ no $(7/8+\varepsilon)$-approx. alg. for Max 3SAT unless $P=NP$.

Best approx ratio for Max SAT is 0.8331.