RECAP

Randomized Algorithm

new instruction : RANDOM(S)

- $X = \mathrm{RANDOM}(S)$ is a random variable with uniform distribution on S:

$$\forall s \in S : \Pr(X = s) = \frac{1}{|S|}$$

- r.v.'s associated to different calls to RANDOM are independent

- Returned value, time (and even correctness) become r.v.'s

- RANDOMIZED APPROXIMATION ALGORITHM: $A^R_\Pi(i)$ returns $s_A \in \mathcal{S}(i)$, s r.v.

- approximation ratio $\rho(|i|)$ of $A^R_\Pi(i) = s_A$ is defined as

$$\rho(|i|) = \max\left\{ \frac{E(c(s_A))}{c(s^*)}, \frac{c(s^*)}{E(c(s_A))} \right\}$$

3-CNF-SAT-related optimization problem:

MAX-3-CNF-SAT

$\mathcal{I} = \langle \varphi(x_1, x_2, \dots, x_n) = \bigwedge_{j=1}^{m} c_j \rangle$

with $C_j = y_j^1 \vee y_j^2 \vee y_j^3$, y_j^k literals of different variables (thus $y_j^{k_1} \neq \bar{y}_j^{k_2}$, $k_1 \neq k_2$)

Determine the maximum number of clauses that can be satisfied by a truth assignment $\vec{b} \in \{0,1\}^n$

DECISION VERSION

3_CNF_MAX_SAT

$\begin{cases} \mathcal{I}: \langle \varphi(x_1, x_2, \dots, x_n) = \bigwedge_{j=1}^{m} c_j, k \leq m \rangle \\ \mathcal{Q}: \exists \vec{b} \in \{0,1\}^n \text{ satisfying} \geq k \text{ clauses?} \end{cases}$

Clearly, $\boxed{\text{3-CNF-SAT} \leq_p \text{3-CNF-MAX-SAT}}$

$\left(\boxed{f(\langle \phi(x_1 \dots x_n) = \bigwedge_{j=1}^{m} c_j \rangle) = \langle \phi, m \rangle} \right)$

(prove correctness of reduction as an EXERCISE)

Simple randomized algorithm for 3-CNF-MAX-SAT: create a random truth assignment $\vec{b}$ and count the number of clauses true under $\vec{b}$:

$$\boxed{\begin{array}{l}
\text{APPROX-3CNF-MAX-SAT}(\phi) \\
\quad * \text{ Let } \phi(x_1,\dots,x_n) = \bigwedge_{j=1}^{m} C_j \ * \\
\quad \textbf{for } i \leftarrow 1 \textbf{ to } n \textbf{ do} \\
\qquad b_i \leftarrow \text{RANDOM}(\{0,1\}); \\
\quad Y \leftarrow 0 \\
\quad \textbf{for } j \leftarrow 1 \textbf{ to } m \textbf{ do} \\
\qquad \textbf{if } (C_j(\vec{b}) = 1) \textbf{ then } Y \leftarrow Y + 1 \\
\quad \textbf{return } Y
\end{array}}$$

RUNNING TIME: $\boxed{\Theta(n+m) = \Theta(k\varphi > 1)}$

APPROXIMATION RATIO Let $\boxed{Y_j(\vec{b})}$ be an indicator variable:

$$\boxed{Y_j = \begin{cases} 0 & \text{if } C_j(\vec{b}) = 0 \\ 1 & \text{otherwise} \end{cases}}$$

We know that $Y_j = 0$ if and only if the three literals of $C_j: y_j^1, y_j^2, y_j^3$ are all false under $\vec{b}$:

e.g. $y_j^1 = x_k$ is false under $\vec{b}$ when $b_k = 0$

and $y_j^1 = \overline{x}_k$ is false under $\vec{b}$ when $b_k = 1$

Since each b_i is set to a random value:

$$\boxed{\Pr\left(y_j^3(\vec{b})=0\right)=\frac{1}{2}.}$$

Since the literals come from different variables and coin flips are independent:

$$\boxed{\Pr(Y_j=0)=\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{8}}$$

$y_j^1=0$ $y_j^2=0$ $y_j^3=0$

Hence $\boxed{\Pr(Y_j=1)=1-\frac{1}{8}=\frac{7}{8}=E[Y_j]}$ $=0\cdot\Pr(0)+1\cdot\Pr(1)=\Pr(1)$

Let $\boxed{Y}$ be the number of satisfied clauses (e.g. the value returned by $A_{3CMS}(\langle\phi\rangle)$). We have:

$$\boxed{Y=\sum_{j=1}^{m}Y_j}$$ (I add 1 only when $C_j(\vec{b})=1$)

Then $\boxed{E[Y]=E\left[\sum_{j=1}^{m}Y_j\right]=\frac{7}{8}m}$

$\parallel$

$E[C(S_{A_{3CMS}})]$

We have: $\boxed{\rho=\frac{C(S^*)}{E[C(S_{A_{3CMS}})]}\le\frac{m}{\frac{7}{8}m}=\frac{8}{7}!}$

REMARK There is also a deterministic $\frac{8}{7}$-approximation algorithm for 3-CNF-MAX-SAT but it is much more complex!

The algorithm is a DERANDOMIZATION of A_{3CNS} yielding a deterministic 8/7-approximation

Håstad in 1997 proved that there cannot exist a ρ-approximation algorithm for 3-CNF-MAX-SAT, with $\rho < \frac{8}{7}$ unless P=NP

MORALE Often the power of randomization allows the development of very simple and efficient algorithms!