

Вероятностное округление для «MAX-SAT»

Н. Н. Кузюрин С. А. Фомин

18 февраля 2011 г.



Вероятностное округление для
«MAX-SAT». Вероятностные алгоритмы,
основанные на округлении
нецелочисленного решения.

Задача «MAX-SAT»

Задача

Максимальная выполнимость/MAX-SAT.

Даны t скобок конъюнктивной нормальной формы (КНФ) с n переменными. Найти значения переменных, максимизирующие число выполненных скобок.

Определение

Литерал — каждое вхождение переменной x_i (или ее отрицания) в скобку. Например, для $(x_1 \vee \bar{x}_2) \wedge (\bar{x}_1 \vee x_3)$ литералами будут $x_1, \bar{x}_2, \bar{x}_1, x_3$.

Определение

Вероятностный приближенный алгоритм A гарантирует точность C , если для всех входов I

$$1 \geq \frac{\mathbf{E} m_A(I)}{m_0(I)} \geq C > 0,$$

где $m_0(I)$ — оптимум, $m_A(I)$ — значение, найденное алгоритмом, и решается задача максимизации.

Теорема

Пусть $x_i = \{0, 1\}$ равновероятно: $p = \frac{1}{2}$. Для любого подмножества скобок $S = \{C_1, \dots, C_m\}$, $|S| = m$

$$\mathbf{E} \sum_{i=1}^m C_i(x) \geq \frac{m}{2}.$$

Доказательство.

Пусть $|C| = k$, тогда

$$P(C = 0) \leq 2^{-k} \Rightarrow P(C = 1) \geq 1 - 2^{-k} \geq \frac{1}{2} \Rightarrow \mathbf{E} C_i(x) \geq \frac{1}{2}.$$

Откуда

$$\mathbf{E} \sum_{i=1}^m C_i(x) \geq \frac{m}{2}.$$



Задача

MAX-SAT(IP)

C_j : скобка.

$z_j = \{0, 1\}$: значение скобки.

C_j^+ : $i : x_i \in C_j$ (положительные литералы в скобке).

C_j^- : $i : \bar{x}_i \in C_j$.

$$\sum_{j=1}^m z_j \rightarrow \max \quad (1)$$

$$\sum_{i \in C_j^+} x_i + \sum_{i \in C_j^-} (1 - x_i) \geq z_j, \quad \forall j.$$

$$x_i, z_j \in \{0, 1\}, \quad \forall i, j$$

Пусть \hat{x}_i, \hat{z}_j — решение линейной релаксации.

$$\begin{aligned} \sum_{j=1}^m z_j &\rightarrow \max & (2) \\ \sum_{i \in C_j^+} x_i + \sum_{i \in C_j^-} (1 - x_i) &\geq z_j, & \forall j. \\ x_i, z_j &\in [0, 1], & \forall i, j \end{aligned}$$

$\sum_{j=1}^m \hat{z}_j$ — верхняя оценка оптимума (2).

Вход: Формулировка задачи 1 в виде (1)
 $\hat{x} \leftarrow$ решения линейной релаксации (2)
for all $i \in \{1..m\}$ **do**
 $x_i \leftarrow 0$
 if $\text{random}(0..1) \leq \hat{x}_i$ **then**
 $x_i \leftarrow 1$ { $x_i \leftarrow 1$ с вероятностью \hat{x}_i }
 end if
end for
Выход: (x_1, \dots, x_m) .

Теорема

Алгоритм обеспечивает приближенное решение (1), со средней точностью $(1 - \frac{1}{e})$.

Лемма

Пусть $|C_j| = k$, $\beta_k = 1 - (1 - 1/k)^k$.

$$P(C_j = 1) \geq \beta_k \hat{z}_j.$$

Доказательство.

Пусть $C_j = x_1 \vee \dots \vee x_k$. Из (2) следует:

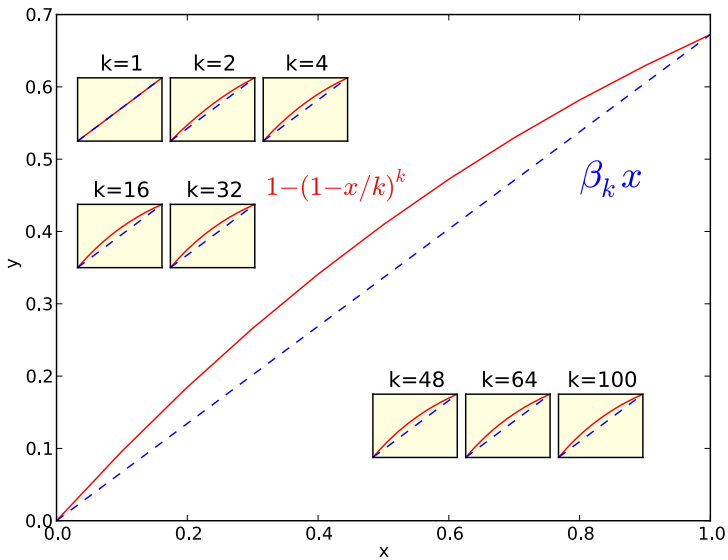
$$\hat{x}_1 + \dots + \hat{x}_k \geq \hat{z}_j.$$

$$P(C_j = 0) = \prod_{i=1}^k (1 - \hat{x}_i) \Rightarrow P(C_j = 1) = 1 - \prod_{i=1}^k (1 - \hat{x}_i).$$

$P(C_j = 1)$ минимально если $\forall i \hat{x}_i = \hat{z}_j/k$.

$f(z) = 1 - (1 - z/k)^k$ — вогнутая функция:

$$\begin{aligned} f(0) &= 0 = \beta_k z \\ f(1) &= 1 - (1 - 1/k)^k = \beta_k z \end{aligned} \Rightarrow f(z) \geq \beta_k z \quad 0 \leq z \leq 1 \quad \forall k \in \mathbb{N}.$$



Теорема

Алгоритм «Линейная релаксация и вероятностное округление» обеспечивает приближенное решение (1), со средней точностью $(1 - \frac{1}{e})$.

Доказательство.

$$\beta_k = 1 - (1 - 1/k)^k \geq 1 - 1/e \quad \forall k \in \mathbb{N}.$$

$$Em_A(I) \geq \sum_k \sum_{C_j \in S^k} \beta_k \hat{z}_j \geq (1 - 1/e) \sum_j \hat{z}_j.$$

$$m_0(I) \leq \sum_j \hat{z}_j.$$



Теорема

Алгоритм «Линейная релаксация и вероятностное округление» обеспечивает приближенное решение (1), со средней точностью $(1 - \frac{1}{e})$.

Доказательство.

$$\beta_k = 1 - (1 - 1/k)^k \geq 1 - 1/e \quad \forall k \in \mathbb{N}.$$

$$Em_A(l) \geq \sum_k \sum_{c_j \in S^k} \beta_k \hat{z}_j \geq (1 - 1/e) \sum_j \hat{z}_j.$$

$$m_0(l) \leq \sum_j \hat{z}_j.$$



$$\left(1 - \frac{1}{e}\right) \approx 0.632120559$$

Алгоритм с вероятностной точностью $\frac{3}{4}$

- 1 n_1 — алгоритм «округление каждой переменной независимо в 0 или 1 с вероятностью $1/2$ »;
- 2 n_2 — алгоритм «Линейная релаксация и вероятностное округление»;
- 3 выбираем $\max(n_1, n_2)$.

Теорема

$$E \max\{n_1, n_2\} \geq \frac{3}{4} \sum_j \hat{z}_j.$$

Доказательство.

S^k : множество скобок, содержащих ровно k литералов. Имеем

$$En_1 = \sum_k \sum_{C_j \in S^k} (1 - 2^{-k}) \geq \sum_k \sum_{C_j \in S^k} (1 - 2^{-k}) \hat{z}_j$$

$$En_2 \geq \sum_k \sum_{C_j \in S^k} \beta_k \hat{z}_j.$$

$$E \max(n_1, n_2) \geq E \frac{n_1 + n_2}{2} \geq \sum_k \sum_{C_j \in S^k} \frac{(1 - 2^{-k}) + \beta_k}{2} \hat{z}_j.$$

$$(1 - 2^{-k}) + \beta_k = 2 - 2^{-k} - \left(1 - \frac{1}{k}\right)^k \geq \frac{3}{2} \quad \forall k \in \mathbb{N}.$$

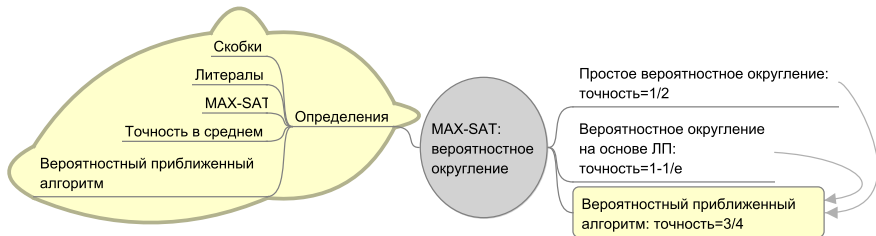
$$E \max(n_1, n_2) \geq \frac{(1 - 2^{-k}) + \beta_k}{2} \hat{z}_j \geq \frac{3}{4} \sum_k \sum_{C_j \in S^k} \hat{z}_j = \frac{3}{4} \sum_j \hat{z}_j.$$

Упражнение

Рассмотрим задачу «MAX-3ESAT», это «MAX-SAT», только в КНФ в каждой скобке ровно три литерала.

Предложите вероятностный алгоритм с точностью $\frac{7}{8}$.

«Карта памяти» лекции



<http://discopal.ispras.ru/>

Вопросы?