Вероятностное округление для «MAX-SAT»

Н. Н. Кузюрин С. А. Фомин

14 апреля 2016 г.

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

Задача «MAX-SAT»

Задача

Максимальная выполнимость/MAX-SAT.
Даны m 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 $(x_1 \vee \overline{x_2}) \wedge (\overline{x_1} \vee x_3)$ литералами будут $x_1, \overline{x_2}, \overline{x_1}, x_3$.*

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 с* $p = \frac{1}{2}$.
Для любого подмножества скобок $S = \{C_1, \ \ldots, C_m\}$, $|S| = m$

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

Доказательство.

Пусть $|C| = k$, тогда

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

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^- : $i : \overline{x_i} \in C_j$.

$$\sum_{j=1}^{m} z_j \to \max \tag{1}$$

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

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

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

$$\sum_{j=1}^{m} z_j \to \max \tag{2}$$
$$\sum_{i\in C_j^+} x_i + \sum_{i\in C_j^-} (1 - x_i) \geq z_j, \quad \forall j.$$
$$x_i,\ z_j \in [0, 1],\ \forall i, j$$

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

Вход: Формулировка задачи 1 «MAX-SAT» в виде (1)

$\hat{x} \leftarrow$ решения линейной релаксации (2)

for all $i \in \{1..m\}$ **do**

$x_i \leftarrow 0$

if $random(0..1) \leq \hat{x}_i$ **then**

$x_i \leftarrow 1$ $\{x_i \leftarrow 1$ с вероятностью $\hat{x}_i\}$

end if

end for

Выход: $(x_1, \ldots, x_m)$.

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.

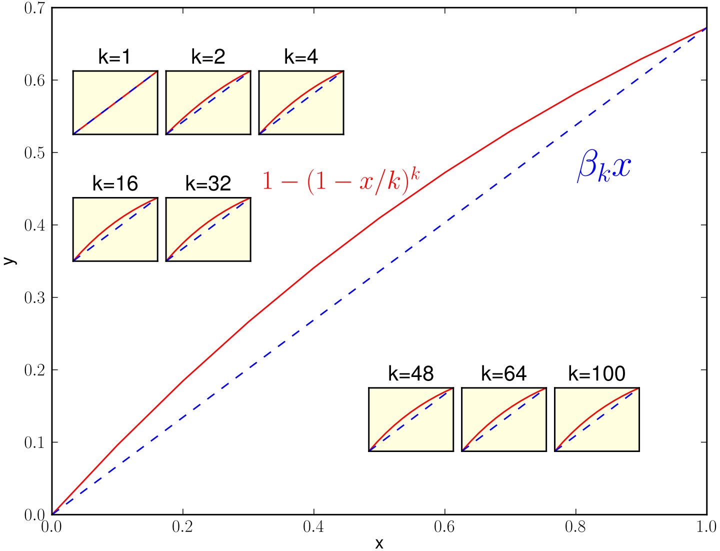
Пусть $C_j = x_1 \vee \ \ldots \vee x_k$. Из (2) следует:

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

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

$P(C_j = 1)$ минимально если $\forall i\, \hat{x}_i = \hat{z}_j/k$.
$f(z) = 1 - (1 - z/k)^k$ — вогнутая функция:

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



k=1
k=2
k=4
k=16
k=32
k=48
k=64
k=100
$1-(1-x/k)^k$
$\beta_k x$
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.0
0.2
0.4
0.6
0.8
1.0
x
y

Теорема

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

Доказательство.

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

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

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

$\square$

Теорема

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

Доказательство.

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

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

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

$\square$

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

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

$$\begin{aligned} En_1 &= \sum_k \sum_{C_j \in S^k} (1 - 2^{-k}) \geq \sum_k \sum_{C_j \in S^k} (1 - 2^{-k}) \hat{z}_j \\ En_2 &\geq \sum_k \sum_{C_j \in S^k} \beta_k \hat{z}_j. \end{aligned}$$

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

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

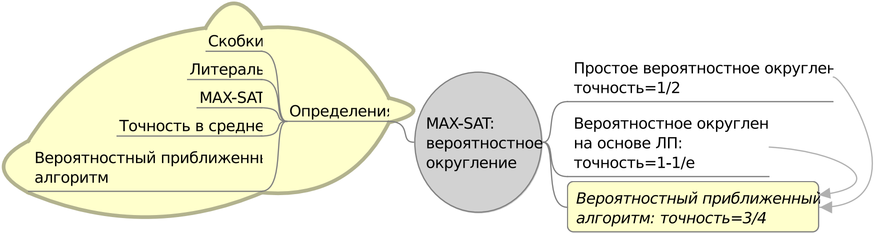
$$E \max(n_1, n_2) \geq \frac{(1 - 2^{-k}) + \beta_k}{2} \hat{z}_j \geq \frac{3}{4} \sum_k \sum_{C_j \in S^k} \hat{z}_j = \frac{3}{4} \sum_j \hat{z}_j.$$

□

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



Интернет поддержка курса

`http://discopal.ispras.ru/`

Вопросы?