УДК 519.854

В.А. МИХАЙЛЮК

ПОДХОД К ОЦЕНКЕ СЛОЖНОСТИ В СРЕДНЕМ ПОСТОПТИМАЛЬНОГО АНАЛИЗА ДИСКРЕТНЫХ ЗАДАЧ ОПТИМИЗАЦИИ

Ключевые слова: *постоптимальный анализ, полиномиальность в среднем относительно* μ.

В теоретических исследованиях информатики основное внимание, как правило, уделялось вопросам анализа сложности проблем в наихудшем случае. Возможно, более естественный (и практический) путь к определению сложности проблем — анализ сложности в среднем (average case complexity). Как известно, многие важные проблемы являются *NP*-трудными, и существует лишь небольшая вероятность их эффективного решения в наихудшем случае. Тогда целесообразны были бы алгоритмы, которые решали бы эти проблемы эффективно в среднем. Это основная мотивация теории сложности в среднем.

При рассмотрении сложности в среднем необходимо задать распределения, которым удовлетворяют экземпляры проблемы. Одна и та же проблема для одних распределений может эффективно решаться в среднем, а для других распределений она труднорешаема.

При решении конкретной оптимизационной задачи с фиксированным набором входных данных обычно получают значительное количество информации, однако лишь часть ее может быть использована для решения именно данной задачи, а остальная часть полученной информации теряется. Это обстоятельство ставит проблему целесообразного использования избыточной информации для решения других оптимизационных задач, в определенном смысле «близких» к исходной.

Проведение постоптимального анализа дискретных оптимизационных задач [1–4] предусматривает решение следующих вопросов.

— Как изменится оптимальное решение конкретной задачи, если некоторым образом изменить значения ее коэффициентов?

— Каким образом использовать информацию, полученную при решении задачи конкретным методом для исследования измененной задачи?

— Какую минимальную дополнительную информацию необходимо накопить при решении исходной задачи в целях эффективного решения измененной задачи?

Наряду со сложностью проблемы в наихудшем случае [5, 6] вызывает интерес применение теории сложности в среднем для постоптимального анализа дискретных задач оптимизации.

1. НЕКОТОРЫЕ ПОНЯТИЯ ТЕОРИИ СЛОЖНОСТИ ВЫЧИСЛЕНИЙ В СРЕДНЕМ

Изложим некоторые необходимые в дальнейшем понятия и определения [7, 8].

Пусть алфавит Σ — упорядоченное, конечное множество символов; Σ^* — множество всех Σ-строк, в дальнейшем $\Sigma = \{0,1\}$. На множестве всех двоичных строк $\{0,1\}^*$ введем стандартный лексикографический порядок: $x < y$ означает, что x предшествует y в лексикографическом порядке, и $y-1$ означает, что непосредственным предшественником является y.

© В.А. Михайлюк, 2011

Определение 1 [8] (функция распределения вероятностей). Функция распределения $\mu : \{0,1\}^* \to [0,1]$ есть неубывающая функция строк на единичный отрезок [0, 1], которая сходится к единице (т.е. $\mu(0) \geq 0$, $\mu(x) \leq \mu(y)$, если $x < y$ и $\lim\limits_{x\to\infty} \mu(x) = 1$). Функцию плотности, связанную с функцией распределения μ, обозначим μ', $\mu'(0) = \mu(0)$ и $\mu'(x) = \mu(x) - \mu(x-1)$ для каждого $x > 0$. Очевидно, что $\mu(x) = \sum\limits_{y \leq x} \mu'(y)$.

Определение 2 [8] (случайная проблема). Случайная проблема распознавания свойств (соответственно случайная проблема поиска) представляет пару (D, μ) (соответственно (S, μ)), где $D : \{0,1\}^* \to \{0,1\}$ (соответственно $S \subseteq \{0,1\}^* \times \times \{0,1\}^*$) и $\mu : \{0,1\}^* \to [0,1]$ есть функция распределения.

В дальнейшем будем вводить понятия для проблем распознавания свойств, подобные понятия для проблем поиска легко модифицируются. Простые распределения обычно связывают с полиномиально вычислимыми (P-вычислимыми) распределениями.

Определение 3 [8] (P-вычислимые распределения). Распределение μ принадлежит классу P-вычислимых, если существует полиномиальная детерминированная машина Тьюринга (ДМТ), которая на входе x выдает двоичное представление $\mu(x)$ (время вычисления полиномиально по $|x|$).

Отсюда следует, что двоичное представление $\mu(x)$ имеет длину, полиномиальную по $|x|$. Если функция распределения μ P-вычислима , то функция плотности μ' вычислима за время, полиномиальное по $|x|$. Обратное неверно, если $P \neq NP$ [7].

Определим класс случайных проблем, который соответствует NP.

Определение 4 [8] (класс DistNP). Случайная проблема распознавания свойств (D, μ) принадлежит классу DistNP, если D есть NP-предикат (принадлежит классу NP) и функция μ P-вычислима.

Определение 5 [8] (полиномиальность в среднем). Функция $f : \{0,1\}^* \to N$ полиномиальна в среднем относительно (распределения) μ, если существует константа $\varepsilon > 0$ такая, что

$$\sum_{x \in \{0,1\}^*} \mu'(x) \cdot \frac{f(x)^{\varepsilon}}{|x|} < \infty .$$

Функция $l(x) = f(x)^{\varepsilon}$ линейна в среднем относительно μ. Функция полиномиальна в среднем, если она оценивается полиномом от функции, которая линейна в среднем.

Определение 6 [7, 8] (класс Average-P). Случайная проблема распознавания свойств (D, μ) разрешима за время APtime, если некоторая ДМТ разрешает D за время, полиномиальное относительно μ в среднем. Average-P представляет класс случайных проблем распознавания свойств, разрешимых за время APtime.

Определение 7 [7]. Функция $f : \Sigma_1^* \to \Sigma_2^*$ вычислима за время APtime по отношению к μ_1 на Σ_1^*, если некоторая ДМТ вычисляет f за время, полиномиальное в среднем относительно μ_1, а $\mu'_2(y) = \sum\limits_{f(x)=y} \mu'_1(x)$ — индуцированная с помощью (f) функция плотности вероятностей на Σ_2^* (с соответствующим распределением μ_2).

Классы Average-P и DistNP являются аналогами соответственно в среднем классов P и NP (в некоторых работах, например [7], эти классы обозначаются AP и RNP).

Основные понятия теории сложности вычислений в среднем введены в рассмотрение Л. Левиным [9]. В работах [7–12] рассмотрены основные достоинства этой теории: общее понятие случайной проблемы, машинно-независимое определение выполнимости алгоритма в среднем, понятия сводимостей различных типов между случайными проблемами, примеры полных проблем для класса NP с достаточно «равномерными» распределениями (DistNP-полные проблемы). Рассматривались также вопросы структурной теории сложности вычислений в среднем. Заслуживает внимания, например, такой результат. Пусть $DTime(f(n))$ (соответственно $NTime(f(n))$) — класс проблем со временем решения, не бóльшим $f(n)$, детерминированной (соответственно недетерминированной) машиной Тьюринга.

Теорема 1 [8]. Если $NTime(2^{O(n)}) \neq DTime(2^{O(n)})$, то DistNP не является подмножеством Average-P.

Замечание 1. Контрапозиция теоремы 1 дает утверждение следующего вида: если $DistNP \subseteq \text{Average-P}$, то $NTime(2^{O(n)}) = DTime(2^{O(n)})$.

Большое значение имеет проблема эквивалентности случайных задач поиска и распознавания свойств в контексте теории сложности в среднем (без такого результата изучение проблем распознавания свойств не отражает структуры проблем поиска). В [8] показано, что случайные проблемы поиска и случайные проблемы распознавания свойств эквивалентны с точки зрения случайных сводимостей Тьюринга.

Пусть $S \subseteq \{0,1\}^* \times \{0,1\}^*$. Уникальной проблемой поиска для S есть проблема для заданного x найти единственное y такое, что $(x, y) \in S$. В [8] показано, что случайные уникальные проблемы поиска и случайные проблемы распознавания свойств эквивалентны с точки зрения детерминированных сводимостей Тьюринга. В дальнейшем будем рассматривать случайные уникальные проблемы поиска. Дискретная оптимизационная задача (S, μ, opt) вытекает из проблемы добавлением оптимизационной функции opt (максимум или минимум), которая соответствует такому единственному y по заданному x, что $(x, y) \in S$.

2. ЭФФЕКТИВНЫЙ ПОСТОПТИМАЛЬНЫЙ АНАЛИЗ В СРЕДНЕМ ДЛЯ ЗАДАЧ ДИСКРЕТНОЙ ОПТИМИЗАЦИИ

Докажем ряд вспомогательных утверждений.

Лемма 1. Если функция $f(x)$ полиномиальна в наихудшем случае (т.е. ограничена полиномом), то она полиномиальна относительно μ в среднем (для произвольного распределения μ).

Доказательство. Предположим, что $|f(x)| \le |x|^n$, $\varepsilon = \dfrac{1}{n+1}$, и используем определение 5 для доказательства, что $f(x) = x^n$ — полиномиальна в среднем относительно μ:

$$\sum_{x \in \{0,1\}^*} \mu'(x) \cdot \frac{f(x)^{\varepsilon}}{|x|} \le \sum_{x \in \{0,1\}^*} \mu'(x) \cdot \frac{|x|^{n/(n+1)}}{|x|} \le$$

$$\le \sum_{x \in \{0,1\}^*} \mu'(x) \cdot |x|^{-1/(n+1)} \le \sum_{x \in \{0,1\}^*} \mu'(x) = 1 < \infty .$$

Лемма 2. Произвольный полиномиальный алгоритм (в наихудшем случае) является полиномиальным в среднем относительно μ алгоритмом (для произвольного распределения μ).

Доказательство. Пусть $A(I)$ — произвольный полиномиальный алгоритм для решения экземпляра I проблемы B. Это означает, что время решения $t_A(I)$ этим алгоритмом экземпляра I есть функция, полиномиальная в обычном смысле. Согласно лемме 1 функция $t_A(I)$ полиномиальна относительно μ в среднем для произвольного распределения μ. Следовательно, A — алгоритм, полиномиальный в среднем относительно μ.

Лемма 3. Если $P = NP$, то $DistNP \subseteq Average\text{-}P$.

Доказательство. Пусть произвольная задача $(S, \mu, opt) \in DistNP$, т.е. S либо NP-трудная, либо NP-полная проблема; μ — P-вычислима. Поскольку $P = NP$, то для S существует полиномиальный по времени алгоритм A (в наихудшем случае), который дает точное решение задачи S. В силу леммы 2 A будет полиномиальным относительно μ в среднем алгоритмом, который дает точное решение задачи S. Согласно определению 6 задача (S, μ, opt) разрешима за время APtime, т.е. $(S, \mu, opt) \in Average\text{-}P$, тем самым лемма доказана.

Рассмотрим произвольную дискретную оптимизационную задачу (S, μ, opt) и пусть $(D, \mu) \in DistNP$ — случайная проблема распознавания свойств, которая ей соответствует. В дальнейшем (S, μ, opt) и (D, μ) будем считать эквивалентными в смысле принадлежности к классам, например, $(D, \mu) \in DistNP$ тогда и только тогда, когда $(S, \mu, opt) \in DistNP$. Будем также считать, что экземпляру I задачи (S, μ, opt) соответствует решение (оптимум) $opt(I)$, если $(I, opt(I)) \in S$. Пусть каждому экземпляру I задачи (S, μ, opt) поставлено в соответствие множество $\{\delta(I)\}$ экземпляров задачи (S, μ, opt). Постоптимальный анализ для пары $(I, \{\delta(I)\})$ даст ответ на вопрос: как при оптимальном решении экземпляра $I(\{opt(I)\})$ найти оптимальное решение экземпляра I' задач с $\{\delta(I)\}$ $(\{opt(I')\})$. Введем понятие эффективного постоптимального сведения в среднем экземпляров задачи $(S, \mu, opt)(\propto^a_{efp})$ ($dom(I)$ — допустимая область задачи с экземпляром I).

Определение 8. Будем считать, что экземпляр $I' \in (S, \mu_1, opt)$ эффективно постоптимально в среднем сводится к экземпляру $I \in (S, \mu, opt)$, когда существует функция f, вычислимая за время APtime относительно распределения μ, что $opt(I') = f(opt(I))$, если $opt(I) \in dom(I')$ (при $opt(I) \notin dom(I')$, функция f — время (число тактов) работы любого полиномиального относительно μ в среднем алгоритма для решения I'), при этом $I \propto^a_{efp} I'$, $\mu'_1(I') = \sum\limits_{f(I)=I'} \mu'(I)$ — индуцированная (f) функция плотности вероятностей на $\{0,1\}^*$ (с соответствующим распределением μ_1).

Определение 9. Для фиксированного экземпляра $I \in (S, \mu, opt)$ обозначим $C(I) = \{I' : I \propto^a_{efp} I'\}$.

Определение 10. Проведение постоптимального анализа для пары $(I, \{\delta(I)\})$ будет эффективным в среднем, если $\{\delta(I)\} \subseteq C(I)$.

Теорема 2. Если DistNP не является подмножеством Average-P, то существует такая пара $(I, \{\delta(I)\})$, для которой проведение постоптимального анализа неэффективно в среднем.

Доказательство. Пусть произвольная задача $(S, \mu, opt) \in DistNP$, $\{P\}$ — множество всех экземпляров задачи (S, μ, opt) и экземпляр $I_1 \in \{P\}$ такой, что для него существует полиномиальный относительно μ в среднем алгоритм, который дает точное решения задачи (такой экземпляр для каждой NP-трудной или NP-полной проблемы (S, μ, opt) существует и зависит от особенностей задачи (S, μ, opt)). Допустим, что для произвольной пары $(I, \{\delta(I)\})$ проведение постоптимального анализа в среднем эффективно.

Положим $I = I_1 \in \{P\}$, $\{\delta(I)\} = \{P\} \setminus I_1$ и рассмотритм пару $(I, \{\delta(I)\})$. Согласно определению 9 $\{P\} \setminus I_1 \subseteq C(I_1)$; отсюда для произвольного $I' \in \{P\} \setminus I_1$ имеем $I_1 \propto^a_{efp} I'$.

В силу выбора I_1 произвольный экземпляр $I \in \{P\}$ можно решить за полиномиальное в среднем время относительно μ, т.е. $(S, \mu, opt) \in Average\text{-}P$. Таким образом, если $(S, \mu, opt) \in DistNP$, то $(S, \mu, opt) \in Average\text{-}P$; следовательно, $DistNP \subseteq Average\text{-}P$, что и доказывает теорему 2.

Замечание 2. Гипотеза, что DistNP не есть подмножеством Average-P, представляется вполне вероятной в силу теоремы 1 и замечания 1 к ней. Согласно лемме 3 (а точнее, контрапозиции к ней) из этой гипотезы вытекает, что $P \neq NP$.

Замечание 3. Теорема 2 показывает сложность проведения постоптимального анализа дискретных задач оптимизации в среднем (что справедливо при изучении сложности постоптимального анализа в наихудшем случае [5]).

Лемма 4 [7]. Пусть заданы функция $f : \Sigma_1^* \to \Sigma_2^*$, μ_1' — функция плотности вероятностей на Σ_1^* и $\mu_2'(y) = \sum\limits_{f(x)=y} \mu_1(x)$ — индуцированная функция плотности вероятностей на Σ_2^*, функция $g : \Sigma_2^* \to \Sigma_3^*$. Если f вычислима за время APtime отностельно μ_1, а g вычислима за время APtime относительно μ_2, то композиция $g \circ f$ (равная $g(f(x))$ для произвольного $x \in \Sigma_1^*$) вычислима за время APtime относительно μ_1.

Рассмотрим произвольную задачу $(S, \mu, opt) \in DistNP$. Пусть I_1, I_2, I_3 — такие экземпляры задачи (S, μ, opt), что $I_1 \propto^a_{efp} I_2 \propto^a_{efp} I_3$. Экземплярам I_j соответствуют распределения вероятностей μ_j $(j = 1, 2, 3)$, и функции f_1 и f_2 (существующие согласно определению 8) таковы, что $opt(I_2) = f_1(opt(I_1))$, $opt(I_3) = f_2(opt(I_2))$, причем f_1 вычислима за время APtime по отношению к μ_1, а f_2 — по отношению к μ_2.

Лемма 5. Пусть $I_1 \propto^a_{efp} I_2 \propto^a_{efp} I_3$, экземплярам I_j соответствуют распределения вероятностей μ_j $(j = 1, 2, 3)$; f_2 вычислима за время APtime по отношению к μ_2, а f_1 — по отношению к μ_1. Тогда композиция $f_2 \circ f_1$ (что соответствует экземплярам I_3) вычислима за время APtime по отношению к μ_1.

Доказательство следует из леммы 4, при этом $\Sigma_1 = \Sigma_2 = \Sigma_3 = \{0, 1\}$, $f = f_1$, $g = f_2$.

Лемма 6. Пусть
$$I_1 \propto^a_{efp} I_2 \propto^a_{efp} \ldots \propto^a_{efp} I_k \propto^a_{efp} I_{k+1} \tag{1}$$
(k — не более чем полином от размерности задачи), и согласно определению 8 экземплярам I_j соответствуют распределения вероятностей μ_j $(j = 1, \ldots$ $\ldots, k, k+1)$, функции f_j таковы, что $opt(I_{j+1}) = f_j(opt(I_j))$, $j = 1, \ldots, k$, при этом f_j вычислима за время APtime по отношению к μ_j $(j = 1, \ldots, k)$. Тогда композиция $f_k \circ f_{k-1} \circ \ldots \circ f_1$ (что соответствует экземплярам I_{k+1}) вычислима за время APtime по отношению к μ_1.

Доказательство проведем индукцией по k в последовательности (1).

При $k = 2$ получим утверждение леммы 5. Допустим, что утверждение имеет место при $k = l$, и покажем, что оно выполняется при $k = l+1$. Имеем последовательность

$$I_1 \propto^a_{efp} I_2 \propto^a_{efp} \ldots \propto^a_{efp} I_l \propto^a_{efp} I_{l+1} \propto^a_{efp} I_{l+2}$$

и композиция $f_l \circ f_{l-1} \circ \ldots \circ f_1$ (что соответствует I_{l+1}) вычислима за время APtime по отношению к μ_1. Согласно определению 8 $opt(I_{l+2}) = f_{l+1}(opt(I_{l+1}))$ и функция f_{l+1} вычислима за время APtime по отношению к μ_{l+1}. Согласно предположению индукции функция $f = f_l \circ f_{l-1} \circ \ldots \circ f_1$ вычислима за время APtime по отношению к μ_1. Применим для $f = f_l \circ f_{l-1} \circ \ldots \circ f_1$ и $g = f_{l+1}$ лемму 4; получим, что композиция $g \circ f = f_{l+1} \circ f_l \circ \ldots \circ f_1$ вычислима за время APtime по отношению к μ_1.

Если D — проблема распознавания свойств и $X \subseteq \Sigma_D^*$, то ограничение $D \mid X$ области D по X есть проблема распознавания свойств с алфавитом $\Sigma(D)$, областью $D \cap X$ и языком $L(D)$. Ограничение случайной проблемы распознавания свойств (D, μ) на множество X экземпляров при $\mu(x) > 0$ есть случайная проблема распознавания свойств $(D \mid X, \mu \mid X)$. Обычно считают, что функция f сводит проблему распознавания свойств D_1 к проблеме распознавания свойств D_2, если $x \in L(D_1)$ для каждого $x \in dom(D_1)$ тогда и только тогда, когда $f(x) \in L(D_2)$.

Определение 11 [7]. 1. Функция f преобразует функцию плотности вероятностей μ'_1 в функцию плотности вероятностей μ'_2, если $\mu'_2(y) = \sum_{f(x)=y} \mu'_1(x)$ для всех y в области μ_2.

2. Функция f преобразует (D_1, μ_1) в (D_2, μ_2), если она сводит $D_1 \mid \{x : \mu_1(x) > 0\}$ к D_2 и преобразует μ'_1 в μ'_2.

Лемма 7 [7]. Допустим, что функция f преобразует (D_1, μ_1) в ограничение (D_2, μ_2), которое принадлежит классу Average-P. Тогда если f вычислима за полиномиальное время или за время, полиномиальное относительно μ_1 в среднем, то (D_1, μ_1) принадлежит Average-*P*.

Рассмотрим дискретную оптимизационную задачу (S, μ, opt), где S — NP-полная проблема поиска, μ — P-вычислима.

Определение 12. Будем считать, что (S, μ, opt) удовлетворяет следующим свойствам *Poly*:

1) для экземпляра I^* существует полиномиальный относительно μ в среднем алгоритм, который дает точное решение задачи;

2) существует последовательность $I^* \propto^a_{efp} I_1 \propto^a_{efp} I_2 \propto^a_{efp} \ldots \propto^a_{efp} I_k \propto^a_{efp} I_{k+1} = I$.

Это возможно тогда и только тогда, когда для произвольного экземпляра $I \in (S, \mu, opt)$ существуют экземпляры $I^*, I_1, \ldots, I_k \in (S, \mu, opt)$ (k — не более чем полином от размерности задачи).

Лемма 8. Задача $(S, \mu, opt) \in Average\text{-}P$ тогда и только тогда, когда (S, μ, opt) удовлетворяет свойствам *Poly*.

Доказательство. 1. Пусть (S, μ, opt) удовлетворяет свойствам *Poly* и $I \in (S, \mu, opt)$ — произвольный экземпляр задачи. Тогда существуют такие экземпляры $I^*, I_1, \ldots, I_k \in (S, \mu, opt)$ (k — не более чем полином от размерности задачи), что

$$I^* \propto^a_{efp} I_1 \propto^a_{efp} I_2 \propto^a_{efp} \ldots \propto^a_{efp} I_k \propto^a_{efp} I_{k+1} = I, \tag{2}$$

экземпляры I_j соответствуют распределениям вероятностей μ_j $(j=1,\ldots$ $\ldots, k, k+1)$, $opt(I_1)=f(opt(I^*))$, $opt(I_{j+1})=f_j(opt(I_j))$, $j=1,\ldots,k$. При этом f_j вычислимы за время APtime по отношению к μ_j $(j=1,\ldots,k)$. Тогда в силу леммы 6 композиция $f_k \circ f_{k-1} \circ \ldots \circ f_1$ (что соответствует экземплярам $I_{k+1}=I$) вычислима за время APtime по отношению к μ_1, т.е. произвольный экземпляр I разрешим за время, полиномиальное относительно μ в среднем; следовательно, $(S, \mu, opt) \in Average\text{-}P$.

2. Пусть $(S, \mu, opt) \in Average\text{-}P$. Покажем, что для произвольного $I \in (S, \mu, opt)$ существуют такие экземпляры I^*, I_1 $(k=1)$, что выполняются условия: 1) для экземпляра I^* имеет место полиномиальный относительно μ в среднем алгоритм, который дает точное решение задачи; 2) существует последовательность $I^* \propto^a_{efp} I_1 \propto^a_{efp} I_2 = I$.

Пусть I — произвольный экземпляр. Выберем произвольное $I^* \neq I$, A — полиномиальный относительно μ в среднем алгоритм, который дает точное решение задачи $I^*(opt(I^*))$. Пусть также $I_1 \neq I^*$, B — полиномиальный относительно μ в среднем алгоритм, который дает точное решение задачи $I_1(opt(I_1))(t_B(I_1)$ — время его работы). Выберем в качестве $f(opt(I^*))$ равное ему $t_B(I_1)$ (независимо от принадлежности $opt(I^*)$ к области $dom(I_1)$). Тогда $opt(I_1)=f(opt(I^*))$ и в силу определения 8 имеем $I^* \propto^a_{efp} I_1$ (экземпляру I_1 соответствует распределение вероятностей μ_1, индуцированное f). Пусть C — полиномиальный относительно μ в среднем алгоритм, который дает точное решение задачи $I(opt(I))$ ($t_C(I)$ — время его работы). Возьмем в качестве $g(opt(I_1))$ равное ему $t_C(I)$ (независимо от принадлежности $opt(I_1)$ к области $dom(I)$). Тогда $opt(I)=g(opt(I_1))$ и в силу определения 8 имеем $I_1 \propto^a_{efp} I$ (экземпляру I соответствует распределение μ_2, индуцированное g), тем самым условие 2 доказано.

Лемма 9. Пусть (S_2, μ_2, opt) — дискретная оптимизационная задача такая, что S_2 является NP-полной, μ_2 — P-вычислимой. Тогда для произвольной $(S_1, \mu_1, opt) \in DistNP$ существует функция f, вычислимая за такое полиномиальное время, что f преобразует (S_1, μ_1, opt) в ограничение (S_2, μ_2, opt). Если $(S_2, \mu_2, opt) \in Average\text{-}P$, то $(S_1, \mu_1, opt) \in Average\text{-}P$.

Доказательство следует из того факта, что S_2 — NP-полна; следовательно, произвольная задача (S_1, μ_1, opt) сводится к (S_2, μ_2, opt) с помощью полиномиальных сводимостей Кука или Карпа [12] (т.е. функция f, полиномиально вычислимая, существует). Далее следует применить лемму 7.

Лемма 10. Если $(S_2, \mu_2, opt) \in Average\text{-}P$, то ввиду произвольности $(S_1, \mu_1, opt) \in DistNP$ следует, что $(S_1, \mu_1, opt) \in Average\text{-}P$.

Доказательство вытекает из леммы 9. Структура утверждения леммы 9 имеет вид $A \to (B \to C)$, а структура утверждения леммы 10 — вид $B \to (A \to C)$. Формулы $A \to (B \to C)$ и $B \to (A \to C)$ равносильны (эквивалентны) как формулы логики высказываний.

Теорема 3. Если (S_2, μ_2, opt) удовлетворяет свойствам *Poly*, то $DistNP \subseteq Average\text{-}P$.

Доказательство вытекает из применимости леммы 8 к лемме 10, а также из того факта, что для произвольной $(S_1, \mu_1, opt) \in DistNP$ следует $(S_1, \mu_1, opt) \in Average\text{-}P$ и $DistNP \subseteq Average\text{-}P$.

Теорема 3 является основной в оценке сложности в среднем постоптимальных процедур задачи (S_2, μ_2, opt). Если к теореме 3 применить принцип контрапозиции, получим следующее утверждение: если DistNP не является подмножеством Average-P, то (S_2, μ_2, opt) не удовлетворяет свойствам *Poly*. Последнее означает, что существует такой экземпляр I_l $(1 \le l \le k)$, что не имеет места сводимость $I_l \propto^a_{efp} I_{l+1}$. Это означает, что для постоптимальной процедуры, которая соответствует сводимости $I_l \propto^a_{efp} I_{l+1}$, не существует полиномиального относительно μ_l в среднем алгоритма.

3. СЛОЖНОСТЬ В СРЕДНЕМ ПОСТОПТИМАЛЬНОГО АНАЛИЗА ЗАДАЧИ О ПОКРЫТИИ МНОЖЕСТВАМИ

Задачу о покрытии множествами будем рассматривать в постановке (задача $\Pi(A, c)$)

$$\min \left\{ f(x) = \sum_{i=1}^{n} c_i x_i \mid x \in Q(A) \right\},$$

$$Q(A) = \left\{ x \in B^n \mid \sum_{j=1}^{n} a_{ij} x_j \ge 1, \ i = 1, \ldots, m \right\},$$

$B^n = \{0, 1\}^n$, $A = \{a_{ij}\}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$; A является (m, n)-булевой матрицей $(m \le n)$; $c_i > 0$, $i = 1, \ldots, n$.

Для булевых (m, n)-матриц $A = \{a_{ij}\}$, $B = \{b_{ij}\}$ введем метрику $\rho(A, B) = \sum_{i,j} | a_{ij} - b_{ij} |$.

Пусть $M = \{1, \ldots, m\}$, $N = \{1, \ldots, n\}$, $K_1 = \{i_1, \ldots, i_k\}$ — некоторая выборка из N объемом k $(1 \le k < n,\ k < m)$. Точка $\alpha = (\alpha_1, \ldots, \alpha_n) \in B^n$ такая, что $\alpha_j = 1$ при $j \in K_1$, $\alpha_j = 0$ при $j \in N \setminus K_1$, а точка $\alpha^i = (\alpha^i_1, \ldots, \alpha^i_n) \in B^n$ такая, что $\alpha^i_i = 1$, $\alpha^i_j = 0$ при $j \ne i$. Опишем класс $\{A_\alpha\}$ булевых (m, n)-матриц $A = \{A_{ij}\}$. Матрица $A \in \{A_\alpha\}$ тогда и только тогда, когда она не содержит одинаковых и нулевых строк, при этом выполняются условия:

1) матрица A имеет подматрицу $A^1 = \begin{pmatrix} \alpha^{i_1} \\ \ldots \\ \alpha^{i_k} \end{pmatrix}$;

2) матрица A имеет подматрицу $A^2 = \{a_{ij}\}$ $(i \in M \setminus K_1, j \in N)$ такую, что для произвольного $i \in M \setminus K_1$ имеет место $\sum_{j \in K_1} a_{ij} \ge 1$, остальные элементы в A^2 произвольны.

Пусть вектор $x^* = (x^*_1, \ldots, x^*_n) \in B^n$ такой, что $x^*_{i_1} = \ldots = x^*_{i_k} = 1$, остальные координаты вектора x^* равны нулю, $A^* \in \{A_\alpha\}$.

Лемма 11 [5]. Вектор x^* — (единственное) оптимальное решение экземпляра I задачи $\Pi(A^*, c)$.

Пусть x^* — оптимальное решение экземпляра I задачи о покрытии $\Pi(A, c)$ и матрица A' такая, что $\rho(A, A') = 1$.

Рассмотрим дискретную оптимизационную задачу (S, μ, opt), которая соответствует $\Pi(A, c)$, и μ — произвольное распределение вероятностей.

Возникает вопрос: можно ли за время, полиномиальное относительно μ в среднем, определить, остается ли x^* оптимальным решением задачи $\Pi(A', c)$? Если не остается, то необходимо найти такое решение. В связи с этим пусть $Find(\Pi(A', c), x^*)$ — процедура, которая определяет оптимальное решение x^* задачи $\Pi(A', c)$, исходя из оптимального решения x^* задачи $\Pi(A, c)$ (результат x^*), которая соответствует сводимости $\propto_{efp}^{a}$. Будем считать, что экземпляр I' $(I' \in \Pi(A', c))$ эффективно постоптимально в среднем относительно μ сводится к $I (I \in \Pi(A, c)) (I \propto_{efp}^{a} I')$, если процедура $Find(\Pi(A', c), x^*)$ может быть выполнена за время, полиномиальное относительно μ в среднем.

Теорема 4. Если DistNP не является подмножеством Average-P, то существуют экземпляр I задачи $\Pi(A', c)$ и такое распределение вероятностей μ, что процедура $Find(\Pi(A', c), x^*)$ не может быть выполнена за время, полиномиальное относительно μ в среднем.

Доказательство. Воспользуемся результатами из [5]. Для задачи (S, μ, opt) применим теорему 3. Пусть A — произвольный (приближенный) полиномиальный относительно μ в среднем алгоритм для решения произвольного $I \in \Pi(A, c)$ и вектор $x^* = (x_1^*, \ldots, x_n^*) \in B^n$ такой, что $x_{i_1}^* = \ldots = x_{i_k}^* = 1$ (остальные координаты вектора x^* равны нулю) есть результат работы алгоритма A на I. Пусть $A^* \in \{A_\alpha\}$ — матрица, соответствующая x^*, в силу леммы 11 алгоритм A для экземпляра I^*(экземпляр $\Pi(A, c)$ с матрицей $A = A^*$) даст точное решение задачи и свойство 1 *Poly* для (S, μ, opt) выполнено.

Пусть $T(A)$ для произвольной булевой (m, n)-матрицы A является преобразованием этой матрицы с заменой произвольного ее одного элемента 0 на 1 или 1 на 0. Будем считать, что $A' = T(A)$, если после применения к A преобразования T получилась матрица A' (очевидно, что $\rho(A, A') = 1$).

Произвольную матрицу A можно получить из матрицы A^*, применяя не более чем $m \cdot n$ преобразований T:

$$A^* \xrightarrow{T} A_1 \xrightarrow{T} A_2 \xrightarrow{T} \ldots \xrightarrow{T} A_k = A, \tag{3}$$

где $A_1 = T(A^*)$, $\rho(A_1, A^*) = 1$, $A_{i+1} = T(A_i)$, $\rho(A_{i+1}, A_i) = 1$, $i = 1, \ldots, k-1$, $k \le m \cdot n \le n^2$.

С применением к последовательности (3) полиномиальной относительно μ процедуры $Find(\Pi(A', c), x^*)$ k раз $(k \le m \cdot n \le n^2)$ получим, что для произвольного экземпляра I задачи $\Pi(A, c)$ существует последовательность I^*, $I_1, I_2, \ldots, I_k, I_{k+1}$ экземпляров, причем

$$I^* \propto_{efp}^{a} I_1 \propto_{efp}^{a} I_2 \propto_{efp}^{a} \ldots \propto_{efp}^{a} I_k \propto_{efp}^{a} I_{k+1},$$

где

$$I_1 \in \Pi(A^*, c), I_2 \in \Pi(A_1, c), I_3 \in \Pi(A_2, c), \ldots$$
$$\ldots, I_k \in \Pi(A_{k-1}, c), I_{k+1} \in \Pi(A_k, c), \; A_k = A.$$

Выполнено свойство 2 *Poly*. Отсюда следует, что $DistNP \subseteq Average\text{-}P$. С применением теоремы 3 (контрапозиции к ней) получим, что существует $I_l \in \Pi(A, c)\,(1 \le l \le k)$ и не выполняется сводимость $I_l \propto^a_{efp} I_{l+1}$. Это значит, что при $I = I_l$, $\mu = \mu_l$ процедура $Find(\Pi(A', c), x^*)$ не может быть выполнена за время, полиномиальное относительно μ в среднем. Теорема доказана.

4. СЛОЖНОСТЬ В СРЕДНЕМ ПОСТОПТИМАЛЬНОГО АНАЛИЗА ЗАДАЧИ О РАНЦЕ

Рассмотрим одномерную булеву задачу о ранце в постановке ($R(a, b, c)$):

$$\max \left\{ f(x) = \sum_{i=1}^{n} c_i x_i \mid x \in Q(a, b) \right\},$$

$$Q(a, b) = \left\{ x \in B^n \mid \sum_{j=1}^{n} a_j x_j \le b \right\}.$$

Будем считать, что все a_j, c_j и b — целые неотрицательные числа: $a = (a_1, \ldots, a_n) \in Z^n$, $c = (c_1, \ldots, c_n) \in Z^n$, $b \in Z^1$; $Z^n = \{x = (x_1, \ldots, x_n) \mid x_i \ge 0,\ x_i$ — целые, $i = 1, \ldots, n\}$. Известно, что задача $R(a, b, c)$ — NP-полна [12].

Для $x = (x_1, \ldots, x_n, x_{n+1}) \in Z^{n+1}$, $y = (y_1, \ldots, y_n, y_{n+1}) \in Z^{n+1}$ введем метрику $\rho(x, y) = \sum_{i=0}^{n+1} |x_i - y_i|$. Пусть $\bar{a} = (a_1, \ldots, a_n, b) \in Z^{n+1}$, а вектор $\bar{a}'$ такой, что $\rho(\bar{a}, \bar{a}') = 1$ (задачу $R(a, b, c)$ будем обозначать также $R(\bar{a}, c)$).

Пусть $N = \{1, 2, \ldots, n\}$, $K_1 = \{i_1, \ldots, i_k\}$ — некоторая выборка из N объема k $(1 \le k \le n)$, вектор $x^* = (x_1^*, \ldots, x_n^*) \in B^n$ такой, что $x_{i_1}^* = \ldots = x_{i_k}^* = 1$, остальные координаты вектора x^* равны нулю. Пусть вектор $\bar{a}^* = (a_1^*, \ldots, a_n^*, b^*)$ такой, что $a_{i_1}^* = \ldots = a_{i_k}^* = 1$, $a_j^* = k + 1$ при $j \in N \setminus K_1$, $b^* = k$.

Лемма 12 [5]. Вектор x^* — (единственное) оптимальное решение экземпляра I задачи $R(\bar{a}^*, c)$.

Рассмотрим дискретную оптимизационную задачу (S, μ, opt), которая соответствует $R(\bar{a}, c)$, и μ — произвольное распределение вероятностей.

Возникает вопрос: если x^* — оптимальное решение экземпляра I задачи $R(\bar{a}, c)$, то можно ли за время, полиномиальное относительно μ в среднем, определить, остается ли x^* оптимальным решением задачи $R(\bar{a}', c)$. Если не остается, то найти такое решение. В связи с этим пусть $Find(R(\bar{a}', c), x^*)$ — процедура, которая определяет оптимальное решение x^* задачи $R(\bar{a}', c)$, исходя из оптимального решения x^* задачи $R(\bar{a}, c)$ (результат x^*), и которая соответствует сводимости $\propto^a_{efp}$. Будем считать, что экземпляр I' $(I' \in R(\bar{a}', c)$ эффективно постоптимально в среднем относительно μ сводится к $I (I \in R(\bar{a}, c)$ $(I \propto^a_{efp} I')$, если процедура $Find(R(\bar{a}', c), x^*)$ может быть выполнена за время, полиномиальное относительно μ в среднем.

Теорема 5. Если DistNP не является подмножеством Average-P, то существуют экземпляр I задачи $R(\bar{a}', c)$ и такое распределение вероятностей μ, что процедура $Find(R(\bar{a}', c), x^*)$ не может быть выполнена за время, полиномиальное относительно μ в среднем.

Доказательство. Воспользуемся результатами из работы [5]. Для задачи (S, μ, opt) применим теорему 3. Пусть А — произвольный (приближенный) полиномиальный относительно μ в среднем алгоритм для решения произвольного $I \in R(\bar{a}, c)$ и вектор $x^* = (x_1^*, \ldots, x_n^*) \in B^n$ такой, что $x_{i_1}^* = \ldots = x_{i_k}^* = 1$ (остальные координаты вектора x^* равны нулю), есть результат работы алгоритма А на I. Пусть вектор $\bar{a}^* = (a_1^*, \ldots, a_n^*, b^*)$ такой, что $a_{i_1}^* = \ldots = a_{i_k}^* = 1$, $a_j^* = k+1$ при $j \in N \setminus K_1$, $b^* = k$. В силу леммы 12 алгоритм А для экземпляра I^* (экземпляр задачи $R(\bar{a}, c)$ при $\bar{a} = \bar{a}^*$) даст точное решение задачи и свойство 1 *Poly* для (S, μ, opt) будет выполнено.

Будем считать, что компоненты вектора $\bar{a}$ ограничены сверху константой c, которая не зависит от размерности задачи n. Будем исходить из вектора $\bar{a}^*$, применяя процедуру $Find(R(\bar{a}', c), x^*)$. Пусть $T(a)$ для произвольного вектора $a \in Z^{n+1}$ является преобразованием вектора a с увеличением или уменьшением одного его компонента на единицу. Будем считать, что $a' = T(a)$, если после применения к a преобразования T получен вектор a' (очевидно, что $\rho(a, a') = 1$).

Произвольный вектор $\bar{a}$ можно получить из вектора $\bar{a}^*$, применяя не более чем $c \cdot (n+1)$ преобразований T:

$$\bar{a}^* \xrightarrow{T} \bar{a}^1 \xrightarrow{T} \ldots \xrightarrow{T} \bar{a}^k = \bar{a}, \tag{4}$$

где $\bar{a}^1 = T(\bar{a}^*)$, $\rho(\bar{a}^1, \bar{a}^*) = 1$; $\bar{a}^{i+1} = T(\bar{a}^i)$, $\rho(\bar{a}^{i+1}, \bar{a}^i) = 1$, $i = 1, \ldots, k-1$; $k \le c \cdot (n+1)$.

Применяя к последовательности (4) полиномиальную относительно μ процедуру $Find(R(\bar{a}', c), x^*)$ k раз ($k \le c \cdot (n+1)$), получим, что для произвольного экземпляра I задачи $R(\bar{a}, c)$ существует последовательность $I^*, I_1, I_2, \ldots, I_k, I_{k+1}$ экземпляров, причем

$$I^* \propto_{efp}^{a} I_1 \propto_{efp}^{a} I_2 \propto_{efp}^{a} \ldots \propto_{efp}^{a} I_k \propto_{efp}^{a} I_{k+1},$$

$$I_1 \in R(\bar{a}^*, c),\ I_2 \in R(\bar{a}^1, c),\ I_3 \in R(\bar{a}^2, c), \ldots$$

$$\ldots,\ I_k \in R(\bar{a}^{k-1}, c),\ I_{k+1} \in R(\bar{a}^k, c),\ \bar{a}^k = \bar{a}.$$

Выполнено свойство 2 *Poly*. Отсюда следует, что $DistNP \subseteq Average\text{-}P$. Применяя теорему 3 (контрапозицию к ней), получим, что существует $I_l \in R(\bar{a}, c)$ $(1 \le l \le k)$, когда не выполняется сводимость $I_l \propto_{efp}^{a} I_{l+1}$, а это означает, что при $I = I_l$, $\mu = \mu_l$ процедура $Find(R(\bar{a}', c), x^*)$ не может быть выполнена за время, полиномиальное относительно μ в среднем. Теорема доказана.

Замечание 4. Из теорем 3 и 5 вытекает сложность проведения постоптимального анализа дискретных задач оптимизации в среднем, что имеет место при анализе сложности и в наихудшем случае постоптимального анализа [5].

Замечание 5. При проведении постоптимального анализа изменение экземпляра I (множество $\{\delta(I)\}$) осуществлялось определенным образом, а не случайно.

Замечание 6. Согласно определению 8 процедура постоптимального анализа эффективна, если для нее существует полиномиальный относительно μ в среднем алгоритм, и неэффективна в противном случае. Пусть $f(x)$ — время работы соответствующей эффективной постоптимальной процедуры, тогда согласно определениям 8 и 5 существует $\varepsilon > 0$, когда математическое ожидание $E\left[\frac{f(x)^{\varepsilon}}{|x|}\right] = O(1)$. Пусть $n = \frac{1}{\varepsilon}$. Найдем вероятность того, что $f(x)$ не более чем полином (в наихудшем случае). Воспользуемся неравенством Маркова [14]: если X — неотрицательная дискретная случайная величина, $h: R \to R$ есть неотрицательная функция и $t > 0$ действительное, то $\Pr[h(X) \ge t] \le \frac{E[h(X)]}{t}$. Имеем

$$\Pr[f(x) \ge k \cdot |x|^{n}] = \Pr\left[\frac{f(x)^{\varepsilon}}{|x|} \ge k^{\varepsilon} \cdot |x|^{\varepsilon n - 1}\right] \le \frac{E\left[\frac{f(x)^{\varepsilon}}{|x|}\right]}{k^{\varepsilon} \cdot |x|^{\varepsilon n - 1}} = \frac{c_1}{k^{\varepsilon}} = O\left(\frac{1}{k^{\varepsilon}}\right).$$

Отсюда $\Pr[f(x) < k \cdot |x|^{n}] \ge 1 - O\left(\frac{1}{k^{\varepsilon}}\right)$, что при больших k близко к единице.

Таким образом, время работы эффективной постоптимальной процедуры не более чем полином с вероятностью, близкой к единице. Значит, время работы неэффективной процедуры экспоненциально с вероятностью, близкой к единице, либо полиномиально на «существенной» части экземпляров (с вероятностью, отличной от единицы), что, по-видимому, и следует считать в качестве труднорешаемости (intractability) постоптимального анализа.

СПИСОК ЛИТЕРАТУРЫ

1. Козерацкая Л.Н., Лебедева Т.Т., Сергиенко И.В. Вопросы устойчивости, параметрический и постоптимальный анализ задач дискретной оптимизации // Кибернетика. — 1983. — № 4. — С. 71–80.
2. Сергиенко И.В., Шило В.П. Задачи дискретной оптимизации. Проблемы, методы решения, исследования. — Киев: Наук. думка, 2003. — 261 с.
3. Geoffrion A.M., Nauss R. Parametric and postoptimality analysis in integer linear programming // Management Sci. — 1977. — **23,** N 5. — P. 453–466.
4. Roodman G.M. Postoptimality analysis in zero one programming by implicit enumeration // Naval Res. Logist. Quart. — 1972. — **19,** N 3. — P. 435–447.
5. Михайлюк В.А. Общий подход к оценке сложности постоптимального анализа дискретных задач оптимизации // Кибернетика и системный анализ. — 2010. — **46,** № 2. — С. 134–141.
6. Михайлюк В.А. Об оценках числовых характеристик сложности постоптимального анализа дискретных задач оптимизации // Там же. — 2010. — **46,** № 5. — С. 136–142.
7. Gurevich Y. Average case completeness // J. Computer and System Sci. — 1991. — **42,** N 3. — P. 346–398.
8. Ben-David S., Chor B., Goldreich O., and Luby M. On the theory of Average Case Complexity // Ibid. — 1992. — **44,** N 2. — P. 193–219.
9. Levin L. Average case complete problems // SIAM J. Comput. — 1986. — **15.** — P. 285–286.
10. Venkatesan R., Levin L. Random instances of a graph coloring problem are hard // Symposium on theory of Computing, ACM. — 1988. — **20.** — P. 217–222.
11. Gurevich Y. Complete and incomplete randomized *NP*-problems // Proc. of the 28th IEEE Symp. On Foundation of Computer Sci. — 1987. — P. 111–117.
12. Levin L.A. One-way functions and pseudo-random generators // Symposium on theory of Computing, ACM. — 1985. — P. 363–375.
13. Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи. — М.: Мир, 1982. — 416 с.
14. Феллер В. Введение в теорию вероятностей и ее приложения. Т. 1. — М.: Мир, 1964. — 498 с.

Поступила 25.11.2010