ЛЕКЦИЯ 16

DISTNP-ПОЛНАЯ ЗАДАЧА. МИРЫ ИМПАЛЬЯЩЦО

1. Distributional problems

Общий вид distributional problems:

$$(A, D),$$

где $A \subset \{0,1\}^*$ — язык, т. е. множество слов, $D = \{D_n\}_{n=1}^{\infty}$ — семейство вероятностных распределений на словах длины n $\{0,1\}^n$.

Рассмотрим **класс distNP**:

$$\text{distNP} = \{(A, D) : A \in \text{NP}, D \text{ считается за полиномиальное время}\}.$$

В отличие от соотношения классов P и NP, для которых справедливо включение P в NP (но не установлено, совпадают ли эти классы):

$$\text{P} \subseteq \text{NP},$$

distP не обязан строго входить в distNP, поскольку в distNP лежат только те пары, для которых D считается за полиномиальное время.

Выясним, вложено ли distNP в distP.

Теорема 30 (Теорема о сводимости) (A, D) сводится к (A', D'), если существует такая функция $f : \{0,1\}^* \to \{0,1\}^*$, что f вычисляется за полиномиальное время и выполняются условия:

1. Корректность: $x \in A \longleftrightarrow f(x) \in A'$;
2. Регулярность по длине: $|f(x)| = p(|x|)$, где $p(|x|)$ — некий полином;
3. Доминирование: $\Pr\{f(D_n) = y\} \leqslant q(n)\Pr\{D'_n = y\}$. $*$

Смысл последнего условия состоит в том, что язык A' долго разрешается на значениях y, которые редко встречаются. Если в соответствие x, на которых язык разрешается быстро, ставить такие y, то, возможно, не получится достичь полиномиального времени.

2. Конструкция distNP-полных задач

Нужно построить distNP-полную задачу и доказать, что она distNP-полная.

Рассмотрим язык u:

$$u = \{(M, x, 1^t) : M(x) = 1, M(x) \text{ работает не больше, чем } t \text{ шагов}\}.$$

Здесь M — недетерминированная машина Тьюринга.

Очевидно, что $u \subseteq \text{NP}$, потому что достаточно предъявить сертификат того, что $M(x) = 1$, после чего остается проверить, что $M(x)$ работает не больше, чем t шагов.

Рассмотрим распределение U:

$U_n : M$ — случайное слово длины $\leqslant \log n$,

T — случайное число от 1 до $n - |M|$,

x — случайное слово длины $n - |M| - t$.

Лемма 2 (Исключение пиков) Для любого семейства распределений $\{D_n\}$ существует такая функция $g : \{0,1\}^* \to \{0,1\}^*$, что:

1. g — инъекция;
2. $|g(x)| \leqslant |x| + 3 + 2\log|x|$;
3. Если x встречается с вероятностью большей или равной 2^{-k}, то $|g(x)| \leqslant k + 3 + 2\log|x|$. *

Док-во: Рассмотрим функцию распределения $\mu_{D_n}(x)$ и функцию $h(x)$ — максимальный общий префикс $\mu_{D_n}(x)$ и $\mu_{D_n}(x-1)$, представленный в двоичной записи.

Возьмём

$$g(x) = \begin{cases} |x|\,0\,x, & \text{если } \Pr\{D_n = x\} < 2^{-n}, \\ 1\,h(x), & \text{если } \Pr\{D_n = x\} \geqslant 2^{-n}. \end{cases}$$

Если x возникает с вероятностью 2^{-k}, то длина общего префикса будет не больше, чем k (потому что $\mu_{D_n}(x)$ и $\mu_{D_n}(x-1)$ суть два двоичных рациональных числа, которые различаются не менее чем на 2^{-k}, а это значит, что хотя бы в k-м знаке возникнет разночтение).

$$\Pr\{D_n = x\} \geqslant 2^{-k} \quad \Rightarrow \quad |h(x)| \leqslant k.$$

Поскольку в слово, которое начинается с 0, не могут перейти два разных слова, то очевидно, что $g = |x|\,0\,x$ есть инъекция.

Пусть $h(x) = h(z)$, $|x| = |z|$. Тогда g есть инъекция в силу монотонности функции $\mu_{D_n}(x)$.

3. Доказательство теоремы о сводимости

Док-во: Пусть машина M рапознает язык L, а M' — язык $L' = \{g(x) : x \in L\}$.

$f(x) = (M', g(x), 1^t)$, где t можно подобрать из соображений регулярности.

Корректность очевидна.

Регулярность можно обеспечить, если выбрать t таким, что

$$t > \text{time}_{M'}(g(x)), \qquad t > 2^{|M'|}.$$

Докажем, что свойство доминирования справедливо. Можно считать, что $M' < \log|f(x)|$.

Оценим вероятность того, что $D_{|f(x)|} = f(x)$. Она не меньше, чем произведение вероятности выбрать M', вероятности выбрать нужные t и вероятности выбрать $g(x)$.

$$\Pr\{D_{|f(x)|} = f(x)\} \geqslant \frac{1}{|f(x)|}\,\frac{1}{|f(x)|}\,\frac{1}{2^{|g(x)|}} \geqslant \frac{1}{|f(x)|^2}\,\frac{1}{2^{|k+3+2\log k|}}.$$

Так как $2^{-|k|} > \frac{1}{2}\Pr\{D_n = x\}$ (поскольку $2^{-|k|} \leqslant \Pr\{D_n = x\} \leqslant 2^{-|k|+1}$), то

$$\Pr\{D_{|f(x)|}\} \geqslant \frac{1}{|f(x)|^2}\,\frac{1}{8k^2}\,\frac{1}{2}\Pr\{D_n = x\}.$$

В итоге:

$$\Pr\{f(D_n) = f(x)\} < 16n^2|f(x)|^2\Pr\{D'_n = f(x)\} = q(n)\Pr\{D'_n = f(x)\}.$$

Свойство доминирования выполняется.

4. Миры Рассела Импальяццо

1. **Алгоритмика.** В этом мире $\mathrm{P} = \mathrm{NP}$. Также здесь быстро доказываются теоремы, но не работает криптография.

2. **Эвристика.** Здесь $\mathrm{P} \neq \mathrm{NP}$, но $\mathrm{distNP} \subset \mathrm{distP}$. В этом мире нет точных алгоритмов, но в среднем они работают хорошо.

3. **Pessiland.** Здесь $\mathrm{P} \neq \mathrm{NP}$, distNP не включается в distP. Также в этом мире не существует односторонних функций (то есть таких функций, которые вычисляются быстро, но прообраз вычисляется медленно).

4. **Миникрипт.** В этом мире существуют односторонние функции, но не существует односторонних функций с секретом (для них вычисление прообраза займет гораздо меньшее время, если знать ключ).

5. **Криптомания.** В этом мире существуют односторонние функции с секретом. Здесь реализуются электронная подпись, электронные выборы, электронные деньги и т. д. Многими считается, что этот мир больше остальных похож на реальный мир.