Kodowanie *

Antoni Kościelski

Wykład 1.

1.1 Kodowanie

Przypuśćmy, że nadawca przesyła odbiorcy informację x. Przyjmujemy, że wtedy odbiorca otrzymuje informację x', która niekoniecznie jest identyczna z x. Chcemy przesyłać informację w taki sposób, aby odbiorca mógł na podstawie tego, co otrzyma ustalić, co nadawca mu przekazywał.

Aby to uzyskać przyjmiemy, że nadawca koduje informację x jako $F(x)$ i zakodowaną przesyła odbiorcy. Odbiorca zaś otrzymuje informację y, która – być może – jest różna od $F(x)$, i na jej podstawie próbuje ustalić $F(x)$, a następnie x.

Pojęcie informacji definiujemy jako słowo nad ustalonym alfabetem (często będzie to zbiór $\{0,1\}$). Wobec tego $F : A^* \to B^*$ dla pewnych alfabetów A i B. Zakładamy, że $F(xy) = F(x)F(y)$ dla dowolnych słów $x, y \in A^*$. Taka funkcja F jest jednoznacznie wyznaczona przez jej obcięcie do zbioru A. Jeżeli $f : A \to B^*$, to symbolem f^* będziemy oznaczać określoną w A^*, przyjmującą wartości w zbiorze B^*, spełniającą $f^*(xy) = f^*(x)f^*(y)$ dla dowolnych słów $x, y \in A^*$ oraz $f^*(a) = f(a)$ dla wszystkich $a \in A$. Funkcję $f : A \to B^*$ nazywamy funkcją kodującą, jeżeli f^* jest funkcją różnowartościową. Oczywiście, że różnowartościowość f^* implikuje różnowartościowość f. Jest też oczywiste, że słowo puste nie jest wartością funkcji kodującej.

Większość interesujących nas własności funkcji kodującej $f : A \to B^*$ wynika z odpowiednich własności zbioru $f(A)$. Będziemy więc mówić, że $K \subseteq B^*$ jest kodem wtedy i tylko wtedy, gdy dla pewnego skończonego zbioru A istnieje funkcja kodująca f przekształcająca A na K.

Kod można też równoważnie zdefiniować jako taki zbiór K, że dla dowolnych słów $x_1, \ldots, x_n, y_1, \ldots, y_m \in K$, warunek $x_1 \ldots x_n = y_1 \ldots y_m$ implikuje, że $n = m$ oraz $x_i = y_i$ dla wszystkich $i = 1, \ldots, n$.

Zbiór K jest kodem przedrostkowym, jeżeli dla dowolnych $x, y \in K$, warunek x jest przedrostkiem y implikuje, że $x = y$. Dowodzi się, że kody przedrostkowe są kodami w sensie wyżej podanej definicji. Zbiór K jest kodem blokowym, jeżeli $K \subseteq B^n$ dla pewnej liczby naturalnej n. Oczywiście, kody blokowe są kodami przedrostkowymi, a więc są też kodami.

*Początkowe wykłady będą prowadzone według książki Foundation of Coding Jiří Adámka

1.2 Nierówność Krafta

Lemat 1.1 *Niech k będzie liczbą liter alfabetu A. Przypuśćmy, że $K = \{s_1, \ldots, s_n\} \subseteq A^*$ jest kodem przedrostkowym. Wtedy zachodzi nierówność Krafta, czyli*

$$\sum_{i=1}^{n} \frac{1}{k^{|\, s_i \,|}} \leqslant 1. \tag{1}$$

Dowód. Niech m będzie liczbą większą od liczb $|\, s_i \,|$ dla $i = 1, \ldots, n$. Będziemy rozważać zbiór A^m. Ma on k^m elementów. Niech $P(s)$ będzie zbiorem tych słów z A^m, które mają przedrostek s. Nietrudno zauważyć, że $P(s)$ ma $k^{m-|\, s \,|}$ elementów. Jest też oczywiste, że jeżeli $P(x) \cap P(y) \neq \emptyset$, to x jest przedrostkiem y lub y jest przedrostkiem x.

Ta ostatnia własność implikuje, że zbiór $\bigcup_{i=1}^{n} P(s_i)$ ma $\sum_{i=1}^{n} k^{m-|\, s_i \,|}$ elementów. Z drugiej strony ten zbiór, jako podzbiór A^m ma najwyżej k^m elementów, Tak więc

$$\sum_{i=1}^{n} k^{m-|\, s_i \,|} \leqslant k^m.$$

Dzieląc otrzymaną nierówność przez k^m otrzymujemy nierówność Krafta. □

Lemat 1.2 *Niech A będzie k literowym alfabetem. Przypuśćmy, że liczby $d_1 \leqslant \ldots \leqslant d_n$ spełniają nierówność*

$$\sum_{i=1}^{n} \frac{1}{k^{d_i}} \leqslant 1.$$

Wtedy istnieje kod przedrostkowy $K = \{s_1, \ldots, s_n\} \subseteq A^*$ taki, że $|\, s_i \,| = d_i$ dla $i = 1, \ldots, n$.□

Twierdzenie 1.3 *Jeżeli $K = \{s_1, \ldots, s_n\} \subseteq A^*$ jest kodem nad k literowym alfabetem A, to spełniona jest nierówność Krafta* (1).

Dowód. Niech

$$c = \sum_{i=1}^{n} \frac{1}{k^{|\, s_i \,|}}.$$

Weźmy też (dowolną) liczbę r i liczbę m niemniejszą od długości wszystkich słów kodowych. Zauważmy, że

$$c^r = (\sum_{i=1}^{n} \frac{1}{k^{|\, s_i \,|}})^r = \sum_{i_1, \ldots, i_r} \frac{1}{k^{|\, s_{i_1} \,|+\ldots+|\, s_{i_r} \,|}} = \sum_{j=1}^{rm} a_j \cdot k^{-j},$$

gdzie a_j jest liczbą tych r elementowych ciągów słów kodowych $s_{i_1}, \ldots, s_{i_r}$ takich, że $|\, s_{i_1} \,| + \ldots + |\, s_{i_r} \,| = j$. Każdy taki ciąg wyznacza słowo $s_{i_1} \ldots s_{i_r}$ długości j. Ponieważ K jest kodem, więc funkcja przyporządkowująca takim ciągom konkatenację ich wyrazów jest wzajemnie jednoznaczna i przekształca zbiór o a_j elementach w zbiór o k^j elementach. Tak więc takich ciągów jest najwyżej k^j. Stąd otrzymujemy, że

$$c^r == \sum_{j=1}^{rm} a_j \cdot k^{-j} \leqslant \sum_{j=1}^{rm} k^j \cdot k^{-j} = r \cdot m.$$

W ten sposób dowiedliśmy, że

$$\frac{c^r}{r} \leqslant m$$

dla wszystkich r. Jest to możliwe tylko wtedy, gdy $c \leqslant 1$. □

1.3 Odległość Hamminga

Przypuśćmy, że mamy dwa k literowe słowa $x = x_1 \dots x_k$ i $y = y_1 \dots y_k$ nad pewnym alfabetem A. Przyjmijmy, że

$$d(x, y) = \#\{i : 1 \leqslant i \leqslant k \wedge x_i \neq y_i\}.$$

Dowodzi się, że funkcja d jest odległością w zbiorze A^k, a więc

1. $d(x, y) \geqslant 0$,
2. $d(x, y) = d(y, x)$,
3. $d(x, y) = 0$ wtedy i tylko wtedy, gdy $x = y$,
4. $d(x, y) \leqslant d(x, z) + d(z, y)$

dla wszystkich $x, y, z \in A^k$.

Funkcję d nazywamy odległością Hamminga. Wartość $d(x, y)$ możemy interpretować jako liczbę błędów w sytuacji, gdy po przesłaniu informacji x jej odbiorca otrzymał informację y.

1.4 Kody wykrywające i korygujące błędy

Przyjmijmy, że

$$S_e(x) = \{y \in A^k : d(x, y) \leqslant e\}.$$

Kod K pozwala wykrywać $\leqslant e$ błędów, jeżeli

$$K \cap \bigcup_{x \in K} (S_e(x) \setminus \{x\}) = \emptyset.$$

Lemat 1.4 *Następujące warunki są równoważne:*

1. *kod K pozwala wykrywać $\leqslant e$ błędów,*
2. $\forall x, y \in K \; x \neq y \Rightarrow d(x, y) > e$,
3. $\forall x, y \in K \; x \neq y \Rightarrow x \notin S_e(y)$.□

Mówimy, że kod K pozwala korygować $\leqslant e$ błędów, jeżeli

$$\forall x, y \in K \; x \neq y \Rightarrow S_e(x) \cap S_e(y) = \emptyset.$$

Lemat 1.5 *Następujące warunki są równoważne:*

1. *kod K pozwala korygować $\leqslant e$ błędów,*
2. $\forall x, y \in K \; x \neq y \Rightarrow d(x, y) > 2 \cdot e$.□