(F1X) Teoria dell'Informazione e della Trasmissione

7 – Disuguaglianza di Kraft-McMillan

Docente: *Nicolò Cesa-Bianchi* versione 31 marzo 2015

I codici di Huffman sono i migliori codici istantanei possibili, nel senso che minimizzano la lunghezza media delle parole di codice. D'altra parte, in certi casi potrebbe essere vantaggioso rinunciare alla proprietà di istantaneità allo scopo di ottenere codici non istantanei (cioè univocamente decodificabili) migliori di quelli di Huffman. Il prossimo risultato mostra invece che il miglior codice univocamente decodificabile non è meglio del codice di Huffman. In particolare, dimostriamo che anche i codici univocamente decodificabili obbediscono alla disuguaglianza di Kraft. Dato che il codice istantaneo ottimo per una sorgente $\langle \mathcal{X}, p\rangle$ di m simboli è definito da

$$\min_{\ell_1,\dots,\ell_m \in \mathbb{N}} \sum_{i=1}^m p_i\,\ell_i \qquad \text{tale che} \quad \sum_{i=1}^m D^{-\ell_i} \le 1$$

ovvero dipende solo dalle lunghezze $\ell_1,\dots,\ell_m$, è evidente che se i codici univocamente decodificabili debbono anch'essi obbedire a Kraft allora l'ottimo dei codici istantanei coincide con l'ottimo dei codici univocamente decodificabili.

Prima di procedere, introduciamo l'estensione k-esima di un codice $c : \mathcal{X} \to \mathcal{D}^+$. Questa è la funzione $C_k : \mathcal{X}^k \to \mathcal{D}^+$ definita come

$$C_k(x_1,\dots,x_k) = c(x_1)\cdots c(x_k)\ .$$

È chiaro che se un codice c è univocamente decodificabile allora la sua k-esima estensione C_k è non singolare per ogni $k \ge 1$. Inoltre, è evidente che $\ell_{C_k}(x_1,\dots,x_k) = \ell_c(x_1) + \cdots + \ell_c(x_k)$ per ogni $x_1,\dots,x_k \in \mathcal{X}$.

Teorema 1 (Disuguaglianza di Kraft-McMillan) $\ell_1,\dots,\ell_m \in \mathbb{N}$ *sono le lunghezze di un codice D-ario univocamente decodificabile per una sorgente di m simboli se e solo se*

$$\sum_{i=1}^m D^{-\ell_i} \le 1\ .$$

Dimostrazione. Prima di tutto osserviamo che se $\ell_1,\dots,\ell_m \in \mathbb{N}$ soddisfano la disuguaglianza di Kraft per un codice D-ario, allora sono le lunghezze di un codice istantaneo —come abbiamo già dimostrato. Dato che un codice istantaneo è univocamente decodificabile, abbiamo dimostrato un verso della doppia implicazione.

Ora dimostriamo che se $c : \mathcal{X} \to \mathcal{D}^+$ è univocamente decodificabile con $|\mathcal{X}| = m$, allora $\ell_c(x_1), \dots, \ell_c(x_m)$ soddisfano la disuguaglianza di Kraft. Per ogni intero $k \geq 1$ possiamo scrivere

$$
\begin{aligned}
\left(\sum_{x \in \mathcal{X}} D^{-\ell_c(x)} \right)^k &= \sum_{x_1 \in \mathcal{X}} \cdots \sum_{x_k \in \mathcal{X}} D^{-\ell_c(x_1)} \times \cdots \times D^{-\ell_c(x_k)} \\
&= \sum_{(x_1,\dots,x_k) \in \mathcal{X}^k} D^{-\left(\ell_c(x_1)+\cdots+\ell_c(x_k)\right)} \\
&= \sum_{(x_1,\dots,x_k) \in \mathcal{X}^k} D^{-\ell_{C_k}(x_1,\dots,x_k)} \ .
\end{aligned}
$$

Ora introduciamo l'insieme $\mathcal{X}_n^k \subseteq \mathcal{X}^k$ di tutti i messaggi che vengono codificati in parole di codice di lunghezza n,

$$
\mathcal{X}_n^k = \Big\{ (x_1, \dots, x_k) \in \mathcal{X}^k \ : \ \ell_{C_k}(x_1, \dots, x_k) = n \Big\} \ .
$$

Possiamo allora scrivere

$$
\sum_{(x_1,\dots,x_k) \in \mathcal{X}^k} D^{-\ell_{C_k}(x_1,\dots,x_k)} = \sum_{n=1}^{k\,\ell_{\max}} \sum_{(x_1,\dots,x_k) \in \mathcal{X}_n^k} D^{-\ell_{C_k}(x_1,\dots,x_k)} = \sum_{n=1}^{k\,\ell_{\max}} \big|\mathcal{X}_n^k\big| D^{-n}
$$

dove $\ell_{\max} = \max\limits_{i=1,\dots,m} \ell_c(x_i)$.

Ora, siccome c è univocamente decodificabile, C_k è non singolare (iniettiva) per ogni $k \geq 1$. Ciò implica che $\mathcal{X}_n^k$ non può contenere più elementi di $\mathcal{D}^n$, altrimenti ci sarebbe una parola di codice in $\mathcal{D}^n$ che codifica due messaggi distinti in $\mathcal{X}_n^k$, violando così l'ipotesi di iniettività. Quindi $\big|\mathcal{X}_n^k\big| \leq \big|\mathcal{D}^n\big| = D^n$. Questo di permette di scrivere

$$
\underbrace{\left(\sum_{x \in \mathcal{X}} D^{-\ell_c(x)} \right)}_{M}^{k} \leq \sum_{n=1}^{k\,\ell_{\max}} \big|\mathcal{X}_n^k\big| D^{-n} \leq \sum_{n=1}^{k\,\ell_{\max}} D^n D^{-n} = k\,\ell_{\max}
$$

dove $M > 0$. La relazione sopra può essere scritta come $M^k \leq k\,\ell_{\max}$. Dato che questa relazione vale per ogni $k \geq 1$, se $M > 1$ ci sarebbe un k_0 tale che per ogni $k \geq k_0$ si avrebbe $M^k > k\,\ell_{\max}$. Quindi dev'essere $M \leq 1$. $\square$