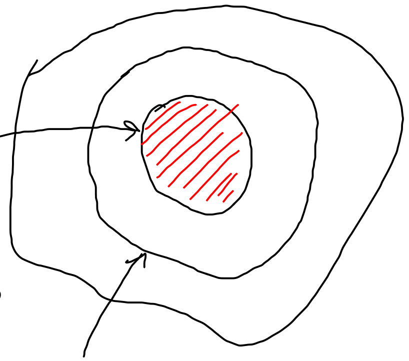
$$C = (\mathcal{M}, \varphi, \mathcal{W}) \qquad \mathcal{M} \subseteq \mathcal{A}^+ , \quad \varphi : \mathcal{M} \to \mathcal{W} , \quad \mathcal{W} \subseteq \mathcal{B}^+$$

$$\mathcal{W} = \varphi(\mathcal{M})$$

Univoca decodificabilità : φ è iniettiva : $\varphi(m) = \varphi(m') \Rightarrow m = m'$.

Se $\mathcal{M} = \mathcal{A}$, si parla di codice simbolico.



Esempio.

$\mathcal{M} = \mathcal{A} = \{a, b, c\}$, $\mathcal{B} = \{0, 1\}$, $\mathcal{W} = \{\varphi(a), \varphi(b), \varphi(c)\} = \{0, 01, 011\}$

Se ricevo $0|0.1|01|0$ l'unica decodifica possibile è abba

Dopo il 1° "0" ancora non posso decodificare con certezza

(ritardo di decodifica).

Il ritardo di decodifica può essere arbitrariamente alto :

Esempio. $\mathcal{M} = \mathcal{A} = \{a, b, c\}$, $\mathcal{W} = \{\varphi(a), \varphi(b), \varphi(c)\} = \{1, 10, 00\}$

$10001 \to bca$

$100001 \to acca$
a c c a

La sequenza $10^n1 = 1\overbrace{0 \ldots 0}^{n \text{ zeri}}1$ decodifica in $\begin{cases} a\,c^{n/2}\,a & \text{se } n \text{ è pari} \\ b\,c^{(n-1)/2}\,a & \text{se } n \text{ è dispari} \end{cases}$ $\to$ Ho un ritardo proporzionale a n

Un codice senza ritardo è detto <u>istantaneo</u>: se ogni messaggio può essere ricostruito immediatamente dopo la ricezione della parola di codice corrispondente.

Def. Una famiglia di parole di codice $W = \{w_1, w_2, \dots, w_T\}$ è detta <u>a prefisso</u> se nessun elemento di W è prefisso di un altro.

Esempio: prefissi telefonici:

$\boxed{\begin{array}{l} 06 \leftarrow \text{roma} \\ 042 \leftarrow \\ 041 \end{array}}$

(Prop. 3.1) Codice è a prefisso $\iff$ è istantaneo.

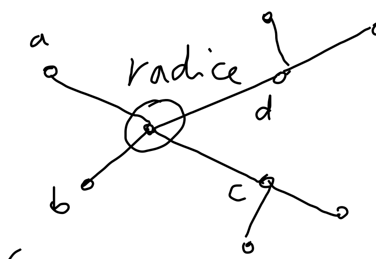
$\Rightarrow$) evidente: non appena una parola di codice, è immediatamente identificabile e decodificabil

$\Leftarrow$) Sia per assurdo: $\alpha, \beta \in W$ tali che $\beta = \alpha\rho$ $\quad (\rho \in B^+)$
(α è prefisso di β)

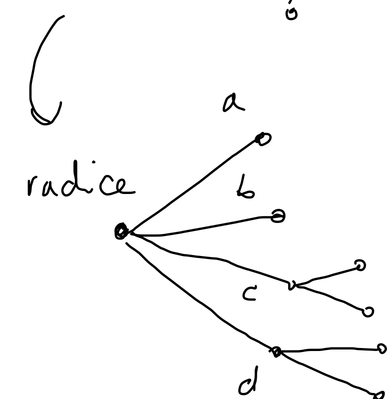
Dopo aver ricevuto la sequenza α non so ancora se decodificare α o aspettare.
Quindi il codice non è istantaneo.

Albero di codice.

Il dizionario W di un codice può essere descritto un albero radicato.
(un grafo aciclico connesso, con un nodo radice evidenziato)



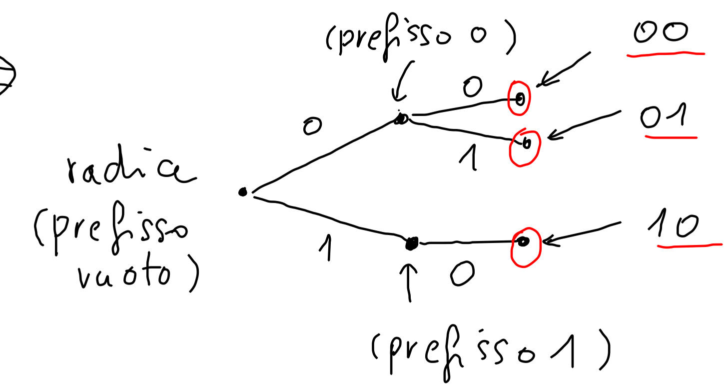
Se l'alfabeto secondario $\mathcal{B}$ ha D simboli,
allora l'albero è un albero D-ario (ogni nodo ha al più D figli)



Il nodo associato alla parola di codice
$\alpha = b_{i_1} b_{i_2} \cdots b_{i_n}$ è il nodo finale del
cammino $b_{i_1}, b_{i_2}, \ldots, b_{i_n}$ che parte dalla radice.

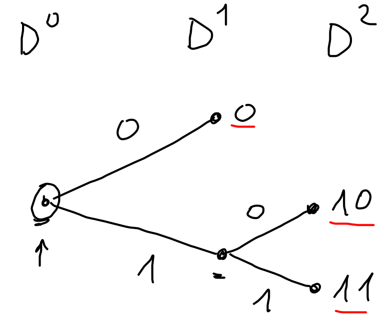
Esempio. $\mathcal{M} = \mathcal{A} = \{a, b, c\}$

$W = \{\varphi(a), \varphi(b), \varphi(c)\} = \{\underline{00}, \underline{01}, \underline{10}\}$ $\Longrightarrow$



Esempio : $\mathcal{A} = \{a, b, c\} = \mathcal{M}$ $\quad W = \{\underset{\varphi(a)}{0}, \underset{\varphi(b)}{10}, \underset{\varphi(c)}{11}\}$

$D = 2 = |\mathcal{B}|$



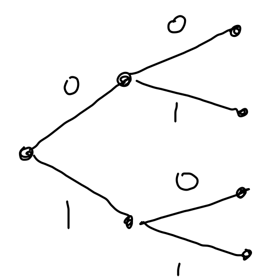
Albero completo ma non zeppo

Albero completo : da ciascun nodo escono esattamente 0 oppure D rami (ogni nodo ha 0 o D figli)

Un Nodo è foglia se ha 0 figli

Albero zeppo : albero completo con D^ℓ nodi di ordine ℓ, per ogni $\ell \le \ell_{max}$

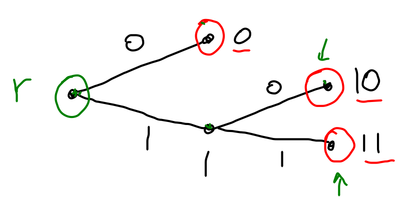
(l'ordine di un nodo è la sua profondità nell'albero)

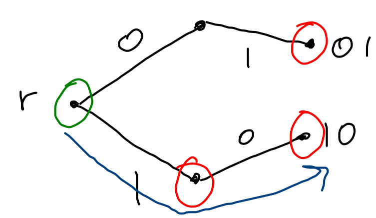


Albero zeppo

Vale la proprietà del prefisso $\iff$ nell'albero di codice si incontra al più 1 parola di codice nel percorso dalla radice ad ogni altro nodo dell'albero.

Esempi. $W = \{0, 10, 11\}$ $\qquad W = \{01, 1, 10\}$





Teo. (3.9) Disuguaglianza di McMillan - Kraft :

Se $l_1, l_2, \dots, l_K$ sono le lunghezze delle parole di codice di un codice u.d., allora

$$(*) \quad \sum_{i=1}^{K} D^{-l_i} \le 1 .$$

(D è la cardinalità di B)
(K è il numero di messaggi, cioè di parole di codice)
$K = |M|$

$$\begin{array}{lcl|l} m_1 & \xrightarrow{\varphi} & \varphi(m_1) = w_1 & l_1 \\ m_2 & \longrightarrow & \varphi(m_2) = w_2 & l_2 \\ \vdots & & \vdots & \vdots \\ m_K & \longrightarrow & \varphi(m_K) = w_K & l_K \end{array}$$

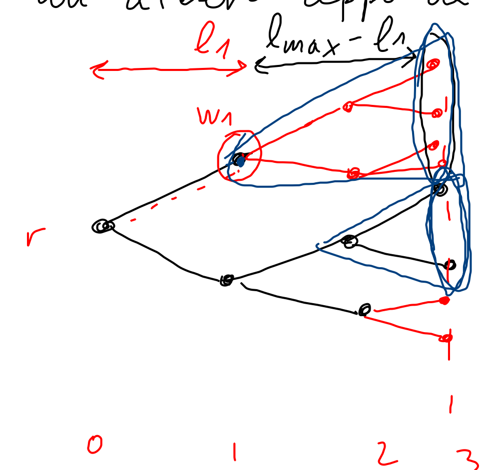
lunghezze

$q \qquad D$

D^q

Dim. Vediamo la dimostrazione nel caso particolare di codici a prefisso.

Consideriamo l'albero D-ario del codice e estendiamolo ad un albero zeppo di altezza $l_{max} = \max(l_1, l_2, \dots, l_K)$.



$D^{l_{max} - l_i}$

Il # di discendenti di livello l_{max} di una parola di codice di lunghezza l_i (che sono anche foglie) è

$D^{l_{max} - l_i}$; tutti questi insiemi di foglie sono disgiunti!

Quindi : $\sum_{i=1}^{K} D^{l_{max}-l_i} \leq D^{l_{max}}$

$\nwarrow$ cardinalità dell'insieme di tutte le foglie

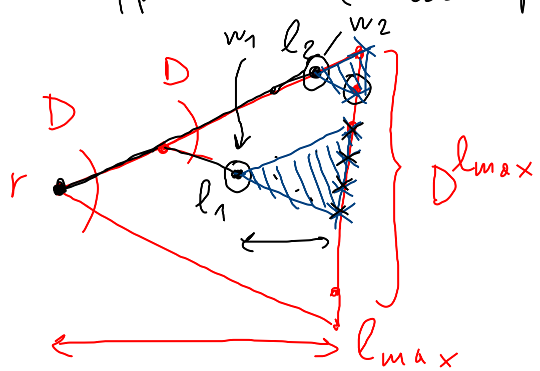
Ma, dividendo per $D^{l_{max}}$, ho :

$\sum_{i=1}^{K} D^{-l_i} \leq 1$. QED.

Teo (3.10). Se gli interi $l_1, l_2, \ldots, l_K$ e D soddisfano la disuguaglianza (*), allora esiste un codice D-ario a prefisso (quindi anche u.d.) in cui ogni parola di codice w_i ha lunghezza l_i.

Dim. Consideriamo l'albero zeppo di profondità $l_{max} = \max(l_1, \ldots, l_K)$.

Supponiamo (senza perdita di generalità) : $l_1 \leq l_2 \leq \ldots \leq l_K$.



Rimuovo $D^{l_{max}-l_1}$ foglie ; $D^{l_{max}-l_1} < D^{l_{max}}$

$\rightarrow$ Ci sono ancora foglie non eliminate

$D^{l_{max}-l_1} + D^{l_{max}-l_2} < D^{l_{max}}$

Posso sempre trovare un percorso di lunghezza l_i per ogni $i = 1, \ldots, K$; il percorso identifica la parola di codice w_i.

Per costruzione, il codice è a prefisso.