Appunti di Teoria dell'informazione

di Vincenzo Russo (vincenzo.russo@neminis.org)

Indice

1. Elementi di Algebra

In questa sezione verrano presentati richiami basilari di algebra, come il concetto di relazione e il concetto di morfismo.

1.1. Relazioni.

DEFINITION 1.1. Una **corrispondenza** R dell'insieme A nell'insieme B è un qualunque sottoinsieme del prodotto cartesiano $A \times B$.

DEFINITION 1.2. Una **relazione** R sull'insieme A è una corrispondenza di A in A, ovvero un qualunque sottoinsieme di $A \times A$. Dati $x, y \in A$, se x è nella relazione R con y scriveremo xRy.

Una relazione R può essere:

- *riflessiva*, se $\forall a \in A$, aRa
- *simmetrica*, se $\forall a, b \in A$, $aRb \implies bRa$
- *antisimmetrica,* se $\forall a, b \in A$, $aRb \wedge bRa \implies a = b$
- *transitiva*, se $\forall a, b, c \in A$, $aRb \wedge bRc \implies aRc$

DEFINITION 1.3. Una relazione che sia al contempo riflessiva, simmetrica e transitiva è detta **relazione di equivalenza**.

DEFINITION 1.4. Una realzione che sia al contempo riflessiva, antisimmetrica e transitiva è invece un **ordinamento parziale**.

1.2. Omomorfismi. Un **omomorfismo** (o **morfismo**) è un'applicazione tra due strutture algebriche dello stesso tipo che conserva le operazioni definite su di esse.

Nel seguito di questo paragrafo considereremo come struttura algebrica d'esempio, un reticolo.

DEFINITION 1.5. Un reticolo $(L, \wedge, \vee)$ è un insieme L con due operazioni binarie definite su di esso: l'operazione di intersezione ($\wedge$) e l'operazione di unione ($\vee$)

DEFINITION 1.6. Siano L e M due reticoli. Un'applicazione $f : L \to M$ è un **omomorfismo** se

- $f(x \wedge y) = f(x) \wedge f(y)$
- $f(x \vee y) = f(x) \vee f(y)$

Se l'omomorfismo è *iniettivo*, si chiama **monomorfismo**; se invece è *suriettivo* allora viene detto **epimorfismo**. Se è biettivo, l'omomorfismo è detto **isomorfismo**. Se $L = M$ allora parliamo di **endomorfismo**. Se l'endomorfismo è biettivo, allora è chiamato **automorfismo**.

Se c'è un isomorfismo da L in M, diremo che L e M sono isomorfi, o anche che L è isomorfo a M e scriveremo $L \cong M$.

In generale, un morfismo da una struttura A a una struttura B ci fornisce una rappresentazione di A attraverso gli elementi di B.

2. Teoria dei semigruppi

Definition 2.1. Un **semigruppo** è un insieme con un'operazione binaria associativa definita su di esso. Dato l'insieme S e l'operazione $\cdot$, scriveremo $(S, \cdot)$ per indicare il semigruppo con S come insieme e $\cdot$ come operazione binaria associativa. E' usuale indicare il semigruppo anche utilizzando soltanto la lettera dell'insieme, laddove l'operazione sia chiara dal contesto o non rilevante.

Definition 2.2. Un semigruppo $(S, \cdot)$ è detto **semigruppo commutativo** se $\forall s, t \in S$, $s \cdot t = t \cdot s$.

Definition 2.3. Sia $(S, \cdot)$ un semigruppo e sia $s \in S$. Allora

(1) s è chiamato **elemento neutro** (o **zero**) di $(S, \cdot)$ se $x \cdot s = s \cdot x = s$, $\forall x \in S$
(2) s è chiamato **elemento identità** di $(S, \cdot)$ se $x \cdot s = s \cdot x = x$, $\forall x \in S$
(3) s è chiamato **idempotente** $s \cdot s = s$

Non tutti i semigruppi hanno lo zero o l'identità; ad ogni modo nessun semigruppo può avere più di uno zero o più di un'identità (ovvero, se l'identità e/o l'elemento neutro esistono, sono unici).

Definition 2.4. Un **monoide** è un semigruppo provvisto dell'elemento identità.

Definition 2.5. Se $(M, \cdot)$ è un monoide e $x \in M$, x è detto **invertibile** se esiste un $y \in M$ tale che $x \cdot y$ e $y \cdot x$ sono uguali all'identità. L'elemento y è detto **inverso** di x e in genere si indica con $y = x^{-1}$ oppure $y = -x$.

Definition 2.6. Un **gruppo** è un monoide in cui tutti gli elementi sono invertibili.

Definition 2.7. Un semigruppo S si dice **cancellativo a sinistra** se $\forall s, t_1, t_2 \in S$, $s \cdot t_1 = s \cdot t_2 \implies t_1 = t_2$

Definition 2.8. Un semigruppo S si dice **cancellativo a destra** se $\forall s, t_1, t_2 \in S$, $t_1 \cdot s = t_2 \cdot s \implies t_1 = t_2$

Definition 2.9. Sia $(S, \cdot)$ un semigruppo. Sia $T \subseteq S$. Se T è chiuso rispetto all'operazione $\cdot$ di S, ovvero:

$$\forall t_1, t_2 \in T \implies t_1 \cdot t_2 \in T$$

allora $(T, \cdot)$ è **sottosemigruppo** di $(S, \cdot)$ e scriveremo $T \leq S$.

Definition 2.10. Sia Γ un insieme di indici e sia S un semigruppo. Sia $(T_\gamma)_{\gamma\in\Gamma}$ una famiglia di sottosemigruppi di S, tale che $\forall\gamma \in \Gamma, T_\gamma \subseteq S$. Allora $T = \bigcap_{\gamma\in\Gamma} T_\gamma$ è un sottosemigruppo di S.

I sottosemigruppi sono dunque chiusi rispetto all'operazione di intersezione.

Definition 2.11. Sia S un semigruppo e sia $X \subseteq S$. L'insieme $< X >= \bigcap_{\gamma\in\Gamma} \{T_\gamma \leq S \mid T_\gamma \supseteq X\}$ [1] è ancora un sottosemigruppo ed è detto *sottosemigruppo generato da* X.

$< X >$ è il più piccolo sottosemigruppo che contiene X.

Definition 2.12. Sia S semigruppo e sia $X \subseteq S$, definiamo $X^+ = X \cup X^2 \cup X^3 \cup ... \cup X^n ... \cup ...$ con $X^n = \{a_1a_2...a_n \mid a_1 \in X, a_2 \in X, ..., a_n \in X\}$

Definition 2.13. Sia S un semigruppo e sia $s \in S$. Allora $s \in X^+ \iff \exists n : s \in X^n$, ovvero tale che $s = x_1x_2x_3...x_n$.

Proposition 2.14. *Sia* S *un semigruppo e sia* $X \subseteq S$. *Allora* $< X >= X^+$.

Definition 2.15. Sia S un semigruppo. $X \subseteq S$ è detto **insieme di generatori** per S se $X^+ =< X >= S$. Se X è un insieme finito, allora S si dice **finitamente generato**.

Definition 2.16. Sia $(M, \cdot)$ un monoide. Sia $T \subseteq M$ e sia 1_M l'identità di M. Se T è chiuso rispetto all'operazione del monoide M e se $1_M \in T$, allora $(T, \cdot)$ è detto sottomonoide di M e scriveremo $T \leq M$.

Definition 2.17. Sia M un monoide e sia $X \subseteq M$. Si definisce sottomonoide generato da X l'insieme $< X >$:

$< X >= \bigcap_{\gamma\in\Gamma} \{T_\gamma \leq M \mid T_\gamma \supseteq X\}$[2]

$< X >= \{1_M\} \cup X \cup X^2 \cup ... \cup X^n \cup ... = \{1_M\} \cup X^+$

Se poniamo $X^0 = \{1_M\}$, possiamo scrivere $< X >= \bigcup_{i\geq 0} X^i = X^0 \cup X^+ = X^*$.

Proprietà degli operatori * *e* $^+$. $\forall X \subseteq S$ con S semigruppo, abbiamo:

- $X \subseteq X^+$ (*estensività*)
- se $X, Y \subseteq S$ e $X \subseteq Y \implies X^+ \subseteq Y^+$ (*isotomia*)
- $(X^+)^+ = X^+$ (*idempotenza*)

Definition 2.18. Siano $(S, \cdot)$ e $(T, \#)$ due semigruppi. Un'applicazione φ da S a T (o viceversa) è detta **morfismo** se $\forall u, v \in S$, $\varphi(u \cdot v) = \varphi(u)\#\varphi(v)$.

[1] Questa intersezione è l'insieme di tutti i sottosemigruppi di S che contengono X.

[2] Come nel caso dei semigruppi, si tratta dell'insieme di tutti i sottomonoidi di M che contengono X.

PROPOSITION 2.19. *Sia $\varphi : S \to T$ un morfismo dal semigruppo S al semigruppo T. Allora vale quanto segue:*

(1) $S^{'} \leq S \implies \varphi(S^{'}) = \left\{\varphi(s) \mid s \in S^{'}\right\} \leq T$ (ovvero: *l'immagine di un sottosemigruppo di S e un sottosemigruppo di T).*
(2) $T^{'} \leq T \implies \varphi^{-1}(T^{'}) = \left\{s \in S \mid \varphi(s) \in T^{'}\right\} \leq S$ (ovvero: *la controimmagine di un sottosemigruppo di T e un sottosemigruppo di S).*

DEFINITION 2.20. Siano $(M_1, \cdot_1, 1_{M_1})$ e $(M_2, \cdot_2, 1_{M_2})$ due monoidi. Un'applicazione φ da M_1 a M_2 (o viceversa) è detta **morfismo** se $\forall s, t \in M_1,\ \varphi(s \cdot_1 t) = \varphi(s) \cdot_2 \varphi(t)$ e $\varphi(1_{M_1}) = 1_{M_2}$.

DEFINITION 2.21. Sia $(S, \cdot)$ un semigruppo. Una relazione θ si dice **compatibile a destra** con $\cdot$ se[3] $\forall s, s_1, s_2 \in S,\ s_1\theta s_2 \implies s_1 s\theta s_2$.

DEFINITION 2.22. Sia $(S, \cdot)$ un semigruppo. Una relazione θ si dice **compatibile a sinistra** con $\cdot$ se $\forall s, s_1, s_2 \in S,\ s_1\theta s_2 \implies ss_1\theta s_2$.

DEFINITION 2.23. Sia $(S, \cdot)$ un semigruppo. Una relazione θ che sia al contempo compatibile a destra e a sinistra si dice semplicemente **compatibile**, ovvero:

$$\forall s_1, s_2, s_3, s_4 \in S, s_1\theta s_2 \wedge s_3\theta s_4 \implies s_1 s_3\theta s_2 s_4$$

DEFINITION 2.24. Sia $(S, \cdot)$ un semigruppo e sia θ una relazione di equivalenza. Si dice che θ è una **relazione di congruenza** se e solo se θ è compatibile con $\cdot$.

Consideriamo ora l'**insieme quoziente** S/θ i cui elementi sono le classi di equivalenza degli elementi di S rispetto alla relazione θ.

Se $s_1 \in S$, allora $\theta(s_1) = \{x \in S \mid s_1\theta x\}$ è la classe di equivalenza di s_1; segue che $S/\theta = \{\theta(s) \mid s \in S\}$.

Definiamo ora l'operazione $*$ come

$$* : S/\theta \times S/\theta \to S/\theta$$

EXAMPLE 2.25. Siano $s_1, s_2 \in S$, con S semigruppo. Allora $\theta(s_1) * \theta(s_2) = \theta(s_1 s_2)$.

L'operazione $*$ è associativa e quindi $(S/\theta, *)$ è un semigruppo, detto **semigruppo quoziente** di $(S, \cdot)$.

DEFINITION 2.26. Sia $\varphi : S \to T$ un morfismo dal semigruppo S al semigruppo T. La relazione di equivalenza θ_φ in S definita come

$$s_1\theta_\varphi s_2 \iff \varphi(s_1) = \varphi(s_2)$$

con $s_1, s_2 \in S$, è la relazione di equivalenza naturalmente indotta da φ su S.

[3] Da questo momento in poi l'operazione $\cdot$ verrà sottintesa, scrivendo ab in luogo di $a \cdot b$.

PROPOSITION 2.27. *Sia $\varphi : S \to T$ un morfismo dal semgriuppo S al semigruppo T. La relazione di equivalenza indotta da φ su S è una congruenza.*

DIMOSTRAZIONE. Siano $s_1, s_2, s_3, s_4 \in S : s_1\theta_\varphi s_2 \wedge s_3\theta_\varphi s_4$. Per la definizione di equivalenza indotta $\varphi(s_1) = \varphi(s_2)$ e $\varphi(s_3) = \varphi(s_4)$, ovvero $\varphi(s_1)\varphi(s_3) = \varphi(s_2)\varphi(s_4)$. Essendo φ un morfismo e per la definizione di equivalenza indotta, si può scrivere $\varphi(s_1s_3) = \varphi(s_2s_4) \iff s_1s_3\theta_\varphi s_2s_4$, che altro non è che la proprietà di compatibilità. Quindi θ_φ è una congruenza. □

DEFINITION 2.28. Sia S un semigruppo e sia θ una congruenza. L'applicazione indotta da θ su S è l'applicazione che associa a ogni elemento di S la sua classe di equivalenza rispetto a θ, ovvero

$$\Psi_\theta : S \to S/\theta$$

definita come $\Psi_\theta(s_1) = \theta(s_1)$, $\forall s_1 \in S$.

Ψ_θ risulta un epimorfismo, detto **epiformismo canonico**.

THEOREM 2.29. (**Teorema di Isomorfismo**) *Sia φ un epimorfismo di S in T, con S e T semigruppi. Consideriamo θ_φ e S/θ_φ, con θ_φ relazione naturalmente indotta da φ. Allora esiste un morfismo biettivo (o isomorfismo) da S/θ_φ in T , ovvero $S/\theta_\varphi \cong T$ (S/θ_φ e T sono isomorfi).*

3. Semigruppi e monoidi liberi

In algebra, una struttura soddisfa la proprietà di Fattorizzazione Unica se ogni elemento può essere scritto in modo unico come "prodotto"[4] di elementi primi.

DEFINITION 3.1. (**Proprietà di fattorizzazione unica**). Sia $(S, \cdot)$ un semi gruppo e sia $X \subseteq S$. X si dice base di S se verifica la proprietà di fattorizzazione unica:
$\forall x_1, x_2, ..., x_h, x_1', x_2', ..., x_k' \in X$ con $h, k \in \mathbb{N}$, se $x_1, ..., x_h = x_1', ..., x_k'$ allora $h = k$ e $\forall i = 1..h, x_i = x_i'$.

La proprietà di fattorizzazione unica garantisce che se un elemento di S può essere fattorizzato come prodotto di elementi di X, tale fattorizzazione è unica.

NOTE 3.2. Se S è un monoide con identità 1, allora $1 \notin X$

DEFINITION 3.3. (**Semigruppo libero**). Sia $(S, \cdot)$ un semigruppo. Esso si dice **libero** se esiste $X \subseteq S$ tale che:

(1) $S = X^+$
(2) X è un base

NOTE 3.4. Se $(S, \cdot)$ è un semigruppo libero, allora esso non può avere l'elemento idendità, ovvero non può essere un monoide.

DEFINITION 3.5. (**Monoide libero**). Sia $(M, \cdot)$ un monoide. Esso si dice **libero** se esiste $X \subseteq M$ tale che:

[4]Il termine prodotto è usato qui in senso lato. Si intende che l'elemento può essere espresso in funzione degli elementi primi tramite l'operazione definita sulla struttura.

(1) $M = X^*$
(2) X è una base

NOTE 3.6. Da quanto già osservato nelle note precedenti, un **monoide libero** non è un **semigruppo libero.**

Monoide delle parole. Supponiamo l'esistenza di un insieme A, tale che $\mid A \mid= r$, con r un valore aribitrario. Costruiamo $Fr(A)$ come l'insieme di tutte le sequenze (*parole*) finite di elementi (*lettere*) di A (*alfabeto*), ovvero

$$Fr(A) = \{w \mid w = a_1 a_2 ... a_n,\, \forall i = 1..n\, a_i \in A\}$$

Se u e v sono parole, allora $u = a_1 a_2 ... a_n$ e $v = b_1 b_2 ... b_m$ con $a_i, b_j \in A$ e $i = 1..n$ e $j = 1..m$.

La parola $uv = a_1 a_2 ... a_n b_1 b_2 ... b_m$ è ottenuta tramite concatenazione delle due parole u e v.

L'operazione di concatenazione è associativa, cioè

$$\forall u, vw \in Fr(A), u(vw) = (uv)w$$

Si scriva dunque tale operazione col simbolo $\cdot$, allora $(Fr(A), \cdot)$ è il semigruppo delle parole e la sua base è A, tale che $Fr(A) = A^+$.

Aggiungendo al semigruppo delle parole la parola vuota ε, otteniamo il monoide delle parole

$$[Fr(A)]^1 = Fr(A) \cup \{\varepsilon\}$$

Nel monoide delle parole l'identità è costituita dalla parola vuota

$$\forall u \in Fr(A)\, u\varepsilon = \varepsilon u = u$$

Ogni monoide libero generato da una base di cardinalità r è isomorfo al monoide delle parole.

PROPOSITION 3.7. *Sia M un monoide (semigruppo) libero e sia X una base per M. Se Y è un insieme di generatori per M, tale che $M = Y^*$ ($M = Y^+$), allora $X \subseteq Y$.*

DIMOSTRAZIONE. Supponiamo per assurdo che $X \nsubseteq Y$. Allora $\exists x \in X \setminus Y$, il che implica $x \in X \wedge x \notin Y$. Poiché $X \subseteq M$, $x \in X \implies x \in M$; essendo Y un insieme di generatori, si può scrivere $x = y_1 y_2 ... y_r$ ed essendo X una base per M, $\forall i = 1..r$, $y_i = x_1^{(i)} x_2^{(i)} ... x_{j_i}^{(i)}$, ovvero possiamo esprimere ogni y_i come prodotto degli elementi della base. Si avrà, dunque, che $x = (x_1^{(1)} ... x_{j_1}^{(1)}) ... (x_1^{(r)} ... x_{j_r}^{(r)})$. X verifica la fattorizzazione unica ed essendo $x \in X$ l'unica possibilità è che siano $r = 1$ e $j_r = 1$; abbiamo dunque $x = x_1^{(1)} = y_1 \in Y \implies x \in Y$, il che ci porta a un assurdo, poiché era stato supposto $x \in X \setminus Y$. □

COROLLARY 3.8. *Una base è un insieme minimo di generatori.*

COROLLARY 3.9. *Sia M un monoide (semigruppo) libero; allora M ha un'unica base.*

PROPOSITION 3.10. (**Caratterizzazione delle basi**). *Sia S un semigruppo e sia X un insieme di generatori per S, tale che $S = X^+$. X è la base di S se e solo se ogni applicazione φ da X in un qualsiasi semigruppo T, $\varphi : X \to T$, si estende a un unico morfismo $\hat{\varphi} : S \to T$.*

DIMOSTRAZIONE. $\Longrightarrow$ Vogliamo dimostrare che preso un qualunque elemento di S e definito $\hat{\varphi}(S)$, $\forall x \in X$ abbiamo che $\hat{\varphi}(x) = \varphi(x)$.

Sia $s \in S$; $s = x_1x_2...x_n$ con $x_i \in X$, $i = 1..n$ e $n \in \mathbb{N}$. Essendo X una base, la fattorizzazione di s è unica. Possiamo scrivere dunque che $\hat{\varphi}(s) = \hat{\varphi}(x_1x_2...x_n)$, e poiché $\hat{\varphi}$ è un morfismo per definizione, $\hat{\varphi}(s) = \hat{\varphi}(x_1)\hat{\varphi}(x_2)...\hat{\varphi}(x_n)$. Ma poiché s è univocamente fattorizzabile tramite gli elementi di X, allora $\hat{\varphi}(s) = \varphi(x_1)\varphi(x_2)...\varphi(x_n)$.

$\Longleftarrow$ Sappiamo che $X^+ = S$. Supponiamo$\mid X \mid= r$ e sia A un insieme tale che $\mid A \mid= r$; poiché A e X hanno la stessa cardinalità, esiste una biezione tra i due insiemi, $\varphi : A \leftrightarrow X$. Associamo ad A il semigruppo di tutte le parole costruite sull'alfabeto A, $F_R(A) = A^+$ e supponiamo inoltre che $\hat{\varphi}$ sia l'estensione di φ tale che $\hat{\varphi} : S \to A^+$ sia suriettiva. Dobbiamo verificare che X soddisfa la proprietà di fattorizzazione unica. Supponiamo per assurdo che X non soddisfi tale proprietà, questo implica che $\exists x_1, x_2, ..., x_h, x_1^{'}, x_2^{'}, ..., x_k^{'} : x_1x_2...x_h = x_1^{'}x_2^{'}...x_k^{'}$ e $h \neq k$ oppure, equivalentemente, $\exists i : x_i \neq x_i^{'}$.

Essendo $\hat{\varphi}$ un morfismo, $\hat{\varphi}(x_1x_2...x_h) = \hat{\varphi}(x_1^{'}x_2^{'}...x_k^{'}) \implies \hat{\varphi}(x_1)\hat{\varphi}(x_2)...\hat{\varphi}(x_h) = \hat{\varphi}(x_1^{'})\hat{\varphi}(x_2^{'})...\hat{\varphi}(x_k^{'})$. Essendo $\hat{\varphi}$ un'estensione di φ per ipotesi, possiamo scrivere $\varphi(x_1)\varphi(x_2)...\varphi(x_h) = \varphi(x_1^{'})\varphi(x_2^{'})...\varphi(x_k^{'})$. Supposto quindi che:

$$\begin{cases} \varphi(x_1) = a_1 \in A \\ \varphi(x_2) = a_2 \in A \\ \qquad \vdots \\ \varphi(x_h) = a_h \in A \end{cases} \text{e} \begin{cases} \varphi(x_1^{'}) = b_1 \in A \\ \varphi(x_2^{'}) = b_2 \in A \\ \qquad \vdots \\ \varphi(x_h^{'}) = b_h \in A \end{cases}$$

si deve avere $a_1a_2...a_h = b_1b_2...b_k$. Poiché A è la base del semigruppo delle parole l'unica possibilità[5] è che $h = k$ e $\forall i = 1..h, a_i = b_i$. Infine, essendo φ una biezione, $\varphi(x_i) = \varphi(x_i^{'}) \implies x_i = x_i^{'}$, ovvero X soddisfa la proprietà di fattorizzazione unica ed è perciò la base. □

PROPOSITION 3.11. *Sia S un semigruppo libero di base X, quindi $X^+ = S$. Sia T un semigruppo libero di base Y, quindi $Y^+ = T$. Supponiamo che $\mid X \mid=\mid Y \mid$, allora $S \cong T$.*

DIMOSTRAZIONE. Essendo $\mid X \mid=\mid Y \mid$ allora esiste una biezione tra X e Y, $\varphi : X \leftrightarrow Y$. Essendo $Y \subseteq T$, allora la biezione tra X e Y è anche un'applicazione da X in T. Essendo X la base di S, per la proposizione 3.10, la biezione si estende a un unico morfismo $\hat{\varphi} : S \to T$. Inoltre, essendo φ una biezione, $\hat{\varphi}$ risulta essere un epimorfismo ($\hat{\varphi}$ è suriettiva). Bisogna dimostrare che è anche iniettiva (ovvero che si tratta di un morfismo biettivo, dunque è un isomorfismo).

Siano $s_1, s_2 \in S : s_1 \neq s_2$, dimostrando che $\hat{\varphi}(s_1) \neq \hat{\varphi}(s_2)$ dimostreremo l'iniettività del morfismo. Supponiamo per assurdo che $\hat{\varphi}(s_1) = \hat{\varphi}(s_2)$; sapendo che $s_1 = x_1x_2...x_h$ e $s_2 = x_1^{'}x_2^{'}...x_k^{'}$, possiamo scrivere $\hat{\varphi}(s_1) = \hat{\varphi}(x_1x_2...x_h) =$

[5]Altrimenti si avrebbe una parola con due fattorizzazioni diverse, che implicherebbe che A non è base, il che è assurdo.

$\hat{\varphi}(x_1^{'}x_2^{'}...x_k^{'}) = \hat{\varphi}(s_2)$. Essendo $\hat{\varphi}$ un morfismo, possiamo riscrivere quest'ultima uguaglianza come $\hat{\varphi}(x_1)\hat{\varphi}(x_2)...\hat{\varphi}(x_h) = \hat{\varphi}(x_1^{'})\hat{\varphi}(x_2^{'})...\hat{\varphi}(x_k^{'})$ e ancora, dato che $\hat{\varphi}$ è un'estensione di φ, possiamo riscriverla come $\varphi(x_1)\varphi(x_2)...\varphi(x_h) = \varphi(x_1^{'})\varphi(x_2^{'})...\varphi(x_k^{'})$. Ma i $\varphi(x_i)$ sono elementi di Y che è una base per T e che quindi soddisfa la proprietà di fattorizzazione unica. Questo comporta che $h = k$ e $\varphi(x_i) = \varphi(x_i^{'})$ $\forall i = 1..h$. Questo implica, essendo φ una biezione, che $x_i = x_i^{'}$ $\forall i = 1..h$, che a sua volta implica che $s_1 = s_2$. Questa è una contraddizione, perché si era supposto $s_1 \neq s_2$. □

Definition 3.12. (**Funzione lunghezza**). Sia M un monoide libero di base A. Ogni elemento $w \in M$, con $w \neq \varepsilon$, si fattorizza in unico modo mediante gli elementi di A. La lunghezza di $w = a_1a_2...a_n$ (con $n \in \mathbb{N}$, $a_i \in A$ $\forall i = 1..n$) è indicato con $\mid w \mid= n$ e corrisponde al numero di fattori della base utilizzati per w.

(*Variante*). Supponiamo di considerare $M \cong [F_r(A)]^1$. Sia $w \in M$ con $w \neq \varepsilon$[6], allora w si può fattorizzare univocamente tramite gli elementi della base:

$w = a_1a_2a_3...a_n$ con $a_i \in A$, $i = 1..n$

$n =\mid w \mid$ è la lunghezza di w.

$\begin{cases} \mid\mid: A \to \mathbb{N} \\ \mid a \mid= 1 \end{cases}$ essendo $M \cong [F_r(A)]^1 = A^*$, si riduce a un unico omomorfismo $\mid\mid: A^* \to \mathbb{N}$.

Proposition 3.13. (**Cancellatività**). *Ogni monoide libero M è cancellativo, cioè*

$\forall u, w_1, w_2 \in M$, se $uw_1 = uw_2 \implies w_1 = w_2$ (cancellatività a sinistra)

oppure

$\forall u, w_1, w_2 \in M$, se $w_1u = w_2u \implies w_1 = w_2$ (cancellatività a destra)

Dimostrazione. Per $u, w_1, w_2 = \varepsilon$ la dimostrazione è banale:

$u = \varepsilon \implies w_1 = w_2$

$w_1 = \varepsilon \implies uw_2 = u \implies \mid u \mid + \mid w_2 \mid=\mid u \mid \implies w_2 = \varepsilon \implies w_1 = w_2$

$w_2 = \varepsilon \implies uw_1 = u \implies \mid u \mid + \mid w_1 \mid=\mid u \mid \implies w_1 = \varepsilon \implies w_1 = w_2$

Supponiamo allora $u, w_1, w_2 \neq \varepsilon$. Sia A la base del monoide M, allora abbiamo $u = a_1a_2...a_n$, $w_1 = b_1b_2...b_r$, $w_2 = c_1c_2...c_s$, con $a_i, b_j, c_k \in A, \forall i = 1..n, \forall j = 1..r, \forall k = 1..s$.

Abbiamo che $uw_1 = uw_2 \implies a_1a_2...a_nb_1b_2...b_r = a_1a_2...a_nc_1c_2...c_s \implies s = r$. Inoltre, essendo A una base, $b_i = c_i$ $\forall i = 1..s$, ovvero $w_1 = w_2$. □

Lemma 3.14. (**Lemma di Levi**). *Sia M monoide libero e sia A la sua base. Siano $w_1, w_2, w_3, w_4 \in M$ tali che $w_1w_2 = w_3w_4$.*

- se $\mid w_1 \mid\geq\mid w_3 \mid$ allora $\exists\xi \in M : \begin{cases} w_1 = w_3\xi \\ \xi w_2 = w_4 \end{cases}$
- se $\mid w_1 \mid<\mid w_3 \mid$ allora $\exists\xi \in M : \begin{cases} w_1\xi = w_3 \\ w_2 = \xi w_4 \end{cases}$

Dimostrazione. Considereremo nella dimostrazione soltanto il primo caso, poiché il secondo si dimostra allo stesso modo.

Sia dunque $\mid w_1 \mid\geq\mid w_3 \mid$ e scriviamo i quattro elementi di M come fattorizzazioni degli elementi della base: $w_1 = a_1a_2...a_r$, $w_2 = b_1b_2...b_s$, $w_3 = c_1c_2...c_p$,

[6] ε è la sequenza (o parola) vuota.

$w_4 = d_1 d_2 ... d_q$, com $a_i, b_j, c_k, d_l \in A$. Allora $r+s = p+q$ e $r \geq p$ il che implica che ci sono p lettere uguali tra w_1 e w_3 , ovvero

$$\begin{cases} a_1 = c_1 \\ a_2 = c_2 \\ \quad \vdots \\ a_p = c_p \end{cases}$$

Essendo M un monoide libero, esso è cancellativo per la proposizione 3.13, e si può scrivere $a_{p+1} a_{p+2} ... a_r b_1 b_2 ... b_s = d_1 d_2 ... d_q$. Ponendo $\xi = a_{p+1} a_{p+2} ... a_r$, scriveremo $\xi w_2 = w_4$. Mentre $w_1 = a_1 a_2 ... a_p a_{p+1} ... a_r = c_1 c_2 ... c_p a_{p+1} ... a_r = w_3 \xi$. $\square$

Proposition 3.15. *Sia M un monoide libero e $N \leq M$, allora N ammette un insieme minimale di generatori, $\dot{N}$, definito come segue:*

$$\dot{N} = (N \setminus 1) \setminus (N \setminus 1)^2$$

Questo implica che $n \in \dot{N} \iff n \neq 1 \wedge n \neq n_1 n_2$, con $n_1, n_2 \in \dot{N}$ e $n_1, n_2 \neq 1$. Ciò significa che gli elementi di $\dot{N}$ sono irriducibili.

Dimostrazione. Per dimostrare questa proposizione procederemo in due fasi. La prima consiste nel dimostrare che $(\dot{N})^* = N$ e la seconda consiste nel dimostrare che $\dot{N}$ è minimale, ovvero che $\forall C : C^* = N \implies \dot{N} \subseteq C$.

Dimostriamo che $(\dot{N})^* \subseteq N$. Supponiamo per assurdo che $(\dot{N})^* \subset N$; questo implica che $\exists n \in N \setminus (\dot{N})^*$.Poiché M è un monoide, tra tutti gli $n \in N \setminus (\dot{N})^*$ possiamo scegliere quella a lunghezza minima, e la indicheremo con $n_0 \in N \setminus (\dot{N})^*$. Quindi $n_0 \notin (\dot{N})^* \implies n_0 \notin \dot{N}$. Sappiamo però che $n_0 \in N$ ed essendo $\dot{N} = (N \setminus 1) \setminus (N \setminus 1)^2$, l'unica possibilità è che $n_0 \in (N \setminus 1)^2$. Questo implica che $\exists n_1, n_2 \in (N \setminus 1) : n_0 = n_1 n_2$; allora abbiamo che $| n_0 | = | n_1 | + | n_2 | \Longrightarrow | n_1 |, | n_2 | < | n_0 |$. Poiché n_0 è stato scelto come elemento a lunghezza minima in $N \setminus (\dot{N})^*$, ciò vuol dire che $n_1, n_2 \in (\dot{N})^*$, ma essendo $n_0 = n_1 n_2 \implies n_0 \in (\dot{N})^*$, che è una contraddizione.

Dimostriamo che $\forall C : C^* = N \implies \dot{N} \subseteq C$. Supponiamo per assurdo che $\exists C : C^* = N$ per il quale si verifica che $\exists n_0 \in \dot{N} \setminus C$. Poiché, come abbiamo visto, $n_0 \in N = C^*$, allora si può scrivere $n_0 = c_1 c_2 ... c_r$ (come prodotto di fattori della base C). Ma $n_0 \in \dot{N}$, quindi è irriducibile, quindi non può essere espresso come prodotto di fattori della base C, quindi $n_0 = c_1 c_2 ... c_r \implies n_0 \in (N \setminus 1)^2$, che è assurdo. $\square$

Definition 3.16. Siano X e Y due insiemi, allora $XY = \{xy \mid x \in X, y \in Y\}$. Ergo, $XX = X^2 = \{xy \mid x \in X, y \in X\}$.

Proposition 3.17. *Sia M un monoide libero e sia $N \leq M$. Sia N generato da un insieme $X \subseteq M$ (ovvero, $N = X^*$) e supponiamo che $1 \notin X$. Allora*

$$\dot{N} = X \setminus X^2 X^*$$

dove $X^2 X^* = \left\{ xy \mid x \in X^2, y \in X^* \right\} = X^2 \cup X^3 \cup ... \cup X^n \cup ...$

Dimostrazione. Sapendo che $N = X^*$, allora abbiamo che $\dot{N} = (N \setminus 1) \setminus (N \setminus 1)^2 = (X^* \setminus 1) \setminus (X^* \setminus 1)^2 = X^+ \setminus (X^+)^2$.

A questo punto ricordiamo che $X^+ = X \cup X^2 \cup ... \cup ... X^n \cup ... = X \cup X^2X^*$. Inoltre $(X^+)^2 = X^+X^+ = \{xy \mid x \in X^+, y \in X^+\} = X^2 \cup X^3 \cup ... \cup X^n \cup ... = X^2X^*$.

Segue quindi che $\dot{N} = (X \cup X^2X^*) \backslash X^2X^*$. Per la proprietà distributiva, $\dot{N} = (X \backslash X^2X^*) \cup (X^2X^* \backslash X^2X^*) = (X \backslash X^2X^*) \cup \oslash = X \backslash X^2X^*$. $\square$

Corollary 3.18. *Un insieme minimale di generatori che sia anche prefisso, è sempre una base, ovvero, sia M un monoide libero e sia $N \leq M$. Sia N generato da un insieme $X \subseteq M$ (ovvero, $N = X^*$) e supponiamo che $1 \notin X$. Allora $X \equiv \dot{N} \iff X \cap X^2X^* = \oslash$.*

Definition 3.19. Sia M un monoide e siano $X, Y \subseteq M$. Possiamo definire

- $X^{-1}Y = \{m \in M \mid Xm \cap Y \neq \oslash\}$ cioè $\exists x \in X, y \in Y : xm = y$ (quoziente o resto sinistro)
- $YX^{-1} = \{m \in M \mid mX \cap Y \neq \oslash\}$ cioè $\exists x \in X, y \in Y : mx = y$ (quoziente o resto destro)

Theorem 3.20. (**Teorema di Schützenberger**). *Sia M un monoide libero e sia N un sottomonoide di M. N è libero $\iff$ $N^{-1}N \cap NN^1 \subseteq N$.*

Dimostrazione. ($\implies$). N libero $\implies N^{-1}N \cap NN^{-1} \subseteq N$. Per dimostrare questo verso dell'implicazione, bisogna provare che

$w \in N^{-1}N \cap NN^{-1} \implies w \in N$, ovvero che $w \in N^{-1}N \wedge w \in NN^{-1} \implies w \in N$, ovvero che

$\exists n_1, n_2, n_3, n_4 \in N : \begin{cases} n_1 w = n_2 \\ w n_3 = n_4 \end{cases} \implies w \in N.$

Se $n_2 = \varepsilon$ o $n_4 = \varepsilon$, allora $\mid n_1 w \mid = \mid w n_3 \mid = 0$ il che implica che $w = \varepsilon \in N$.

Supponiamo invece che

$n_1 = a_1 a_2 ... a_h$ con $a_i \in \dot{N}, \forall i = 1..h$

$n_2 = b_1 b_2 ... b_k$ con $b_i \in \dot{N}, \forall i = 1..k$

$n_3 = c_1 c_2 ... c_r$ con $c_i \in \dot{N}, \forall i = 1..r$

$n_4 = d_1 d_2 ... d_s$ con $d_i \in \dot{N}, \forall i = 1..s$

Allora possiamo scrivere

$$\begin{cases} n_1 w = n_2 \\ w n_3 = n_4 \end{cases} = \begin{cases} (a_1 ... a_h) w = (b_1 ... b_k) \\ w(c_1 ... c_r) = (d_1 ... d_s) \end{cases}$$

Moltiplichiamo ambo i membri della prima equazione per $(c_1 ... c_r)$ ottenendo

$$(a_1 ... a_h) w (c_1 ... c_r) = (b_1 ... b_k)(c_1 ... c_r)$$

Dalla seconda equazione vediamo che $w(c_1 ... c_r) = (d_1 ... d_s)$, quindi possiamo riscrivere la prima equazione come segue

$$(a_1 ... a_h)(d_1 ... d_s) = (b_1 ... b_k)(c_1 ... c_r)$$

Da quest'ultimo sviluppo si evince che $h + s = k + r$.

Possiamo avere quindi tre casi.

Se $h < k$, allora $a_1 = b_1 ... a_h = b_h$, ovvero i primi h elementi di n_1 sono gli stessi dei primi h di n_2 e possiamo sviluppare ancora l'equazione scrivendo

$$(a_1 ... a_h)(d_1 ... d_s) = (b_1 ... b_h)(b_{h+1} ... b_k)(c_1 ... c_r)$$

dato che $a_1 = b_1 ... a_h = b_h$ allora, per la cancellatività si ha

$$(d_1...d_s) = (b_{h+1}...b_k)(c_1...c_r)$$

e poiché $w(c_1...c_r) = (d_1...d_s)$

$$w(c_1...c_r) = (b_{h+1}...b_k)(c_1...c_r) \implies w = (b_{h+1}...b_k) \implies w \in N$$

Se invece $h = k$, questo implica $w = \varepsilon \in N$.

Se, infine, $h > k$, ci troveremmo in un assurdo, poiché i primi k elementi di n_1 sono gli stessi dei primi k di n_2 e possiamo sviluppare ancora l'equazione scrivendo

$$(a_1...a_k)(a_{k+1}...a_h)(d_1...d_s) = (b_1...b_k)(c_1...c_r)$$

dato che $a_1 = b_1...a_k = b_k$ allora, per la cancellatività si ha

$$(a_{k+1}...a_h)(d_1...d_s) = (c_1...c_r)$$

Poiché $(d_1...d_s) = w(c_1...c_r)$ allora

$$(a_{k+1}...a_h)w(c_1...c_r) = (c_1...c_r)$$

e, di nuovo per la cancellatività, si ha

$$(a_{k+1}...a_h)w = \varepsilon$$

il che implica che

$$(a_1...a_h)w = \varepsilon$$

che implicherebbe $a_i = \varepsilon \ \forall i = 1..h$, il che è impossibile visto che $a_i \in \dot{N}$ e $\varepsilon \notin \dot{N}$.

($\Longleftarrow$). $N^{-1}N \cap NN^{-1} \subseteq N \implies N$ libero , ovvero che N ha la base, ovvero che è soddisfatta la proprietà di fattorizzazione unica.

Presi $a_i \in \dot{N} \ \forall i = 1..h$ e $b_i \in \dot{N} \ \forall i = 1..k$, dimostreremo che $(a_i)_{i=1..h} = (b_i)_{i=1..k} \implies h = k$ e $\forall i = 1..h, a_i = b_i$.

Supponiamo $| \ b_1 \ |\geq| \ a_1 \ |$. Per li Lemma di Levi, abbiamo

$\begin{cases} b_1 = a_1 z \\ a_2...a_h = z(b_2...b_k) \end{cases} \implies \begin{cases} z \in N^{-1}N \\ z \in NN^{-1} \end{cases} \implies z \in N^{-1}N \cap NN^{-1} \implies z \in N$

Ma sappiamo che $b_1 = a_1 z$, $b_1 \in \dot{N}$, $a_1 \in \dot{N}$, il che significa che è possibile scrivere b_1 come prodotto di elementi di $\dot{N}$, il che è assurdo perché essendo b_1 un elemento di $\dot{N}$, è irriducibile. Per questo l'unica possibilità è che $z = \varepsilon$, che implica che $b_1 = a_1$.

Reiterando il ragionamento tutti i b_i e a_i, si possono verificare due casi:

$h = k$: in questo caso avremmo dimostrato l'asserto, poiché avremmo mostrato che $a_i = b_i$ per ogni $i = 1..h$

$h > k$ oppure $k > h$: se così fosse, ci si troverebbe, a un certo punto, con $a_{k+1}...a_h = \varepsilon$ (oppure $a_{h+1}...a_k = \varepsilon$), cioè $a_i = \varepsilon$ per ogni $i = 1..h$ (o $i = 1..k$), il che è impossibile in quanto $a_i \in \dot{N}$ e $\varepsilon \notin \dot{N}$.

Quindi l'unica possibilità è che $h = k$, che conclude la nostra dimostrazione. □

THEOREM 3.21. (**Teorema di Cohn**). *Sia M un monoide libero e sia $N \leq M$ (sottomonoide),*

N è libero $\iff \forall m \in M, n \in N$: $\begin{cases} nm \in N \\ mn \in N \end{cases} \implies m \in N$

Il teorema di Cohn è equivalente al teorema di Schützenberger.

Applicazioni del teorema di Schützenberger.

PROPOSITION 3.22. *Sia M monoide libero e siano $N_\gamma \leq M$ con $\gamma \in \Gamma$ sottomonoidi liberi di M; comunque presi gli N_γ abbiamo che $N = \bigcap_{\gamma\in\Gamma} N_\gamma$ è un sottomonoinde libero di M.*

DIMOSTRAZIONE. Per dimostrare che $N \leq M$ basta utilizzare il teorema di Cohn e dimostrare quindi che

$\forall m \in M, \forall n_1n_2n_3n_4 \in N$: $\begin{cases} n_1m = n_2 \\ mn_3 = n_4 \end{cases} \implies m \in N$

Per definizione, $N = \bigcap_{\gamma\in\Gamma} N_\gamma$, quindi $n \in N \implies n \in N_\gamma \forall\gamma \in \Gamma$. Poiché ogni N_γ è libero vale che

$\forall m \in M, \forall n_1n_2n_3n_4 \in N$: $\begin{cases} n_1m = n_2 \\ mn_3 = n_4 \end{cases} \implies m \in N_\gamma$

per ogni N_γ, quindi $m \in N_\gamma \forall\gamma \in \Gamma$ e cioè $m \in \bigcap_{\gamma\in\Gamma} N_\gamma = N$ - c.v.d. □

Quest'ultima proposizione afferma in sostanza che i sottomonoidi liberi sono chiusi rispetto all'operazione di intersezione.

DEFINITION 3.23. Sia S un semigruppo e X una parte di S ($X \subseteq S$). Se $X^{-1}X \cap XX^{-1} \subseteq X$, allora X è detta ***parte liberabile*** di S (detto ***semigruppo stabile***).

In seguito a questa definizione, il teorema di Schützenberger può essere riformulato come segue:

THEOREM. (**Teorema di Schützenberger**). *Sia M un monoide libero e sia N un sottomonoide di M. N è libero $\iff$ N è parte liberabile di M.*

DEFINITION 3.24. Sia M un monoide libero e sia N un sottomonoide. N si dice ***unitario a sinistra*** se $N^{-1}N \subseteq N$. N si dice ***unitario a destra*** se $NN^{-1} \subseteq N$.

DEFINITION 3.25. Sia M un monoide libero e sia N un sottomonoide. Se N è unitario sia a sinistra che a destra, N si dice ***biunitario***.

PROPOSITION 3.26. *Sia M un monoide libero e $N \leq M$ e $M = A^*$. Se N è unitario a sinistra allora N è libero e la sua base $X = (N \setminus 1) \setminus (N \setminus 1)^2$ è un insieme prefisso ($X \cap XA^+ = \oslash$).*

DIMOSTRAZIONE. Per ipotesi N è unitario a sinistra, ovvero $N^{-1}N \subseteq N$. E' banale che $N^{-1}N \cap NN^{-1} \subseteq N^{-1}N$. Segue che $N^{-1}N \cap NN^{-1} \subseteq N$, ovvero N è parte liberabile di M. Per il teorema di Schützenberger, N è libero. Ora resta da dimostrare che X base di N è un insieme prefisso.

Supponiamo per assurdo che X non sia prefisso, questo significa $\exists x_1, x_2 \in X$: $x_1 = x_2z$. Se così fosse, $z \in N^{-1}N \subseteq N \implies z \in N$ e cioè x_1 sarebbe il prodotto di elementi di N, il che è assurdo in quanto X è costituito da elementi irriducibili, quindi X è prefisso. □

4. Teoria dei Codici

$$S \to C \to Canale \to D \to R$$

S è detta *sorgente*, C *codificatore*, D *decodificatore*, R *ricevitore.*

Definition 4.1. La sorgente S è descritta mediante un alfabeto $Y = \{y_1, y_2, ..., y_h\}$ detto alfabeto sorgente. Y^* è un monoide libero su Y e ogni $n \in Y^*$ è detto *messaggio sorgente* (o *messaggio in chiaro*), poiché Y^* costituisce l'insieme delle parole costruite sull'alfabeto Y, che è anche la base del monoide libero Y^*.

Definition 4.2. Sia A un insieme detto alfabeto dei codici e sia A^* il monoide libero sull'alfabeto dei codici, allora si definisce *codifica sequenziale* un morfismo iniettivo (un monomorfismo) dall'insieme dei messaggi in chiaro all'insieme dei messaggi in codice, $\varphi : Y^* \to A^*$ tale che

$\forall w_1, w_2 \in Y^* \implies \varphi(w_1 w_2) = \varphi(w_1)\varphi(w_2)$ (proprietà di morfismo)

$\forall w_1, w_2 \in Y^*, w_1 \neq w_2 \implies \varphi(w_1) \neq \varphi(w_2)$ (proprietà di iniettività)

Definition 4.3. Un codice su A è la base di un sottomonoide libero di A^*.

Proposition 4.4. *Siano Y^* e A^* monoidi liberi, sia Y la base di Y^* e sia $\varphi : Y^* \to A^*$ un monomorfismo. Allora $\varphi(Y) = X$ è un codice su A (o codice per A^*).*

Dimostrazione. Supponiamo per assurdo che X non sia un codice; allora $\exists h, k \in \mathbb{N} : x_1, x_2, ..., x_h, x_1^{'}, x_2^{'}, ..., x_k^{'} \in X$ e $x_1 x_2 ... x_h = x_1^{'} x_2^{'} ... x_k^{'}$ e $h \neq k$ (oppure $\exists i : x_i \neq x_i^{'}$). Essendo φ un monomorfismo ed essendo $\varphi(Y) = X$ tra X e Y esiste una biezione. Dunque $\exists y_1, ..., y_h, y_1^{'}, ..., y_k^{'}$ tali che

$$\begin{cases} \varphi(y_1) = x_1 \\ \varphi(y_2) = x_2 \\ \quad\vdots \\ \varphi(y_h) = x_h \end{cases} \text{e} \begin{cases} \varphi(y_1^{'}) = x_1^{'} \\ \varphi(y_2^{'}) = x_2^{'} \\ \quad\vdots \\ \varphi(y_k^{'}) = x_k^{'} \end{cases}$$

quindi

$\varphi(y_1)...\varphi(y_h) = \varphi(y_1^{'})...\varphi(y_k^{'}) \implies \varphi(y_1...y_h) = \varphi(y_1^{'}...y_k^{'})$; poiché φ è iniettiva, allora $y_1...y_h = y_1^{'}...y_k^{'}$ ed essendo Y una base, h deve essere uguale a k, che è assurdo poiché si è supposto $h \neq k$. □

Proposition 4.5. *Viceversa, se X è un codice su A e Y è un insieme che ha la stessa cardinalità di X, allora ogni biezione $\varphi : Y \leftrightarrow X$ si estende a un monomorfismo di Y^* in A^*.*

Dimostrazione. Essendo X un codice su A, esso è una base per un sottomonoide libero di A^* quindi per la proposizione 3.10 a pagina 10,$\varphi : Y \leftrightarrow X$ si estende a un unico morfismo $\hat{\varphi}^{'} : Y^* \to X^*$ e quindi anche a un morfismo $\hat{\varphi} : Y^* \to A^*$ essendo $X^* \leq A^*$.

Resta da dimostrare l'iniettività di $\hat{\varphi}$, ovvero che si tratta di un monomorfismo, come recita la tesi, e cioè che presi $w_1, w_2 \in Y^* : w_1 = y_1...y_h, w_2 = y_1^{'}...y_k^{'}$ e $\hat{\varphi}(w_1) = \hat{\varphi}(w_2)$ allora deve verificarsi che $w_1 = w_2$.

Per le proprietà dei morfismi e per il fatto che $\hat{\varphi}$ è un'estensione di φ, abbiamo
$\hat{\varphi}(w_1) = \hat{\varphi}(y_1...y_h) = \hat{\varphi}(y_1)...\hat{\varphi}(y_h) = \varphi(y_1)...\varphi(y_h)$
$\hat{\varphi}(w_2) = \hat{\varphi}(y_1^{'}...y_k^{'}) = \hat{\varphi}(y_1^{'})...\hat{\varphi}(y_k^{'}) = \varphi(y_1^{'})...\varphi(y_k^{'})$
Essendo per ipotesi φ una biezione, $\exists x_1, ..., x_h, x_1^{'}, ..., x_k^{'} \in X$ tali che

$$\begin{cases} \varphi(y_1) = x_1 \\ \varphi(y_2) = x_2 \\ \vdots \\ \varphi(y_h) = x_h \end{cases} \text{e} \begin{cases} \varphi(y_1^{'}) = x_1^{'} \\ \varphi(y_2^{'}) = x_2^{'} \\ \vdots \\ \varphi(y_k^{'}) = x_k^{'} \end{cases}$$

Avendo supposto $\hat{\varphi}(w_1) = \hat{\varphi}(w_2)$ abbiamo che $x_1...x_h = x_1^{'}...x_k^{'}$ ed essendo X un codice su A[7], deve essere $h = k$ e $x_i = x_i^{'}$ $\forall i = 1..h$; ma essendo φ una biezione, allora anche $y_i = y_i^{'}$ $\forall i = 1..h$ e quindi $w_1 = w_2$ - c.v.d. $\square$

Resti destri e resti sinistri. Sia M un monoide e sia $X \subseteq M$.

Il resto destro di primo ordine è $R_1 = X^{-1}X \setminus \{\varepsilon\}$; quello di secondo è dato da $R_2 = X^{-1}R_1 \cup R_1^{-1}X$.

Si giunge quindi al resto di ordine n, dato da $R_n = X^{-1}R_{n-1} \cup R_{n-1}^{-1}X$.

Discorso simile per i resti sinistri.

Il resto sinistro di primo ordine è $L_1 = XX^{-1} \setminus \{\varepsilon\}$. Quello di ordine n è $L_n = XL_{n-1}^{-1} \cup L_{n-1}X^{-1}$.

Teorema di Sardinas e Patterson. Il teorema di Sardinas e Patterson ci conferisce la facoltà di sapere se un particolare insieme X è codice in base ai suoi resti destri[8]. La dimostrazione di tale teorema è coadiuvata da due lemmi che nel seguito saranno presentati prima della dimostrazione del teorema stesso.

LEMMA 4.6. *Sia* $X \subseteq A^+$*; allora* $\forall n > 0, R_n \cap X \neq \oslash \iff \varepsilon \in R_{n+1}$.

DIMOSTRAZIONE. ($\implies$) $R_n \cap X \neq \oslash \implies \exists x : x \in X \wedge x \in R_n$.

Poiché $x \in R_n$, per definizione di R_{n+1} sottraiamo ad x, x stesso ottenendo ε. Pertanto $\varepsilon \in R_{n+1}$.

($\impliedby$) $\varepsilon \in R_{n+1} \implies_{per\, definizione} \varepsilon \in X^{-1}R_n$ oppure $\varepsilon \in R_n^{-1}X$.

Nel primo caso abbiamo: $\varepsilon \in X^{-1}R_n \implies \exists x \in X \wedge \exists r_n \in R_n : x\varepsilon = r_n \implies r_n = x \implies R_n \cap X = \oslash$

Nel secondo caso: $\varepsilon \in R_n^{-1}X \implies \exists x \in X \wedge \exists r_n \in R_n : r_n\varepsilon = x \implies r_n = x \implies R_n \cap X = \oslash$ $\square$

LEMMA 4.7. *Sia* $X \subseteq A^+$. $\forall n > 0$ $\forall k = 1..n$ *si ha che* $\varepsilon \in R_n \iff \exists r_k \in R_k \exists i, j \in \mathbb{N} : i + j + k = n \wedge r_k X^i \cap X^j \neq \oslash$[9]

DIMOSTRAZIONE. La dimostrazione avviene per induzione discendente su k.

Passo base ($k = n$).

($\implies$) Per ipotesi $\varepsilon \in R_n$ ed essendo $k = n$ gli unici valori di i e j tali che $i + j + k = n$ sono $i = 0$ e $j = 0$; esisterà quindi un $r_k \in R_k \equiv R_n$ tale che $r_k X^0 \cap X^0 = r_k \{\varepsilon\} \cap \{\varepsilon\}$. Si può quindi scegliere $r_k = \varepsilon$ ottenendo così $\{\varepsilon\} \cap \{\varepsilon\} \neq \oslash$ e dimostrando pertanto l'asserto.

[7] Quindi X è una base e quindi soddisfa la proprietà di fattorizzazione unica.

[8] Equivalentemente possono essere usati i resti sinistri, ma nel seguito si farà sempre uso dei resti destri.

[9] $r_k X^i \cap X^j$ equivale a dire che dato $r_k \in R_k$ esistono $x_1, x_2, ..., x_i, x_1^{'}, x_2^{'}, ..., x_j^{'} \in X$ tali che $r_k x_1 ... x_i = x_1^{'} ... x_j^{'}$.

($\Longleftarrow$) Per ipotesi $\exists r_k \in R_k \equiv R_n$ e $\exists i, j \in \mathbb{N} : i + j + k = n$ il che implica, essendo $k = n$, che $i = j = 0$. Avremo, quindi, $r_k X^0 \cap X^0 \implies r_k \{\varepsilon\} \cap \{\varepsilon\} \neq \oslash \implies r_k = \varepsilon \implies \varepsilon \in R_n$, come dovevasi dimostrare.

Passo di induzione.

($\implies$) Supponiamo vera la tesi per $k + 1$, ovvero la nostra ipotesi induttiva è

$$\varepsilon \in R_n \implies \exists r_{k+1} \in R_{k+1} \exists i, j \in \mathbb{N} : i + j + k + 1 = n \wedge r_{k+1} X^i \cap X^j \neq \oslash$$

Per definizione $R_{k+1} = X^{-1} R_k \cup R_k^{-1} X$; ciò vuol dire che $r_{k+1} \in X^{-1} R_k$ oppure $r_{k+1} \in R_k^{-1} X$ e pertanto

$$\exists x \in X \exists r_k \in R_k : \begin{cases} x r_{k+1} = r_k \ (Caso\, 1) \\ \quad oppure \\ r_k r_{k+1} = x \ (Caso\, 2) \end{cases}$$

Caso 1. Per ipotesi induttiva vale

$$r_{k+1} x_1 x_2 ... x_i = x_1^{'} x_2^{'} ... x_j^{'}$$

ricordando che $r_{k+1} X^i \cap X^j \neq \oslash$. Moltiplicando a sinistra ambo i membri per x si ottiene

$$x r_{k+1} x_1 x_2 ... x_i = x x_1^{'} x_2^{'} ... x_j^{'}$$

Poiché siamo nel caso 1, $x r_{k+1} = r_k$ e quindi

$$r_k x_1 x_2 ... x_i = x x_1^{'} x_2^{'} ... x_j^{'}$$

Questo significa che $r_k X^i \cap X^{j^{'}} \neq \oslash$ con $i + j^{'} + k = n$, $r_k \in R_k$, $i, j^{'} \in \mathbb{N}$ e $j^{'} = j + 1$. Dall'arbitrarietà di i e j segue l'asserto.

Caso 2. Per ipotesi induttiva abbiamo ancora una volta

$$r_{k+1} x_1 x_2 ... x_i = x_1^{'} x_2^{'} ... x_j^{'}$$

Moltiplicando a sinistra ambo i membri per r_k otteniamo

$$r_k r_{k+1} x_1 x_2 ... x_i = r_k x_1^{'} x_2^{'} ... x_j^{'}$$

Trovandoci nel caso 2, $r_k r_{k+1} = x$ e quindi abbiamo

$$x x_1 x_2 ... x_i = r_k x_1^{'} x_2^{'} ... x_j^{'}$$

Questo significa che $X^{i^{'}} \cap r_k X^j \neq \oslash$ con $i^{'} + 1 + j + k = n$, $r_k \in R_k$, $i,^{'} j \in \mathbb{N}$ e $i^{'} = i + 1$. Dall'arbitrarietà di i e j segue l'asserto.

($\Longleftarrow$) Supponiamo vera, ancora una volta, la tesi per $k + 1$, ovvero la nostra ipotesi induttiva è

$$\exists r_{k+1} \in R_{k+1} \exists i, j \in \mathbb{N} : i + j + k + 1 = n \wedge r_{k+1} X^i \cap X^j \neq \oslash \implies \varepsilon \in R_n$$

Per ipotesi invece abbiamo che $\exists r_k \in R_k \exists i^{'}, j^{'} \in \mathbb{N} : r_k X^{i^{'}} \cap X^{j^{'}} \neq \oslash$ con $i^{'} + j^{'} + k = n$. Dunque abbiamo

$$r_k x_1 x_2 ... x_{i^{'}} = x x_1^{'} x_2^{'} ... x_{j^{'}}^{'}$$

Applichiamo ora il Lemma di Levi e consideriamo i due casi:

1) $| x_1^{'} | \geq | r_k |$
2) $| r_k | \geq | x_1^{'} |$
Caso 1: per il lemma di Levi abbiamo
$$\begin{cases} x_1^{'} = r_k p \\ x_1 ... x_{i'} = p x_2^{'} ... x_{j'}^{'} \end{cases} \implies p \in R_{k+1}$$
Poniamo allora $r_{k+1} = p$ e otteniamo $r_{k+1} \in X^{j^{'}-1} \cap X^{i^{'}} \neq \oslash$ con $k + 1 + i^{'} + j^{'} - 1 = n \implies \varepsilon \in R_n$.

Per ipotesi induttiva e per l'arbitrarietà di i e j, si è giunti alla dimostrazione dell'asserto.

Caso 2: procedendo analogamente al caso 1, si ha che
$$\begin{cases} r_k = x_1^{'} p \\ p x_1 ... x_i = x_2^{'} ... x_j^{'} \end{cases} \implies p \in R_{k+1}$$
Poniamo allora $r_{k+1} = p$ e otteniamo $r_{k+1} \in X^{i^{'}} \cap X^{j^{'}-1} \neq \oslash$ con $k + 1 + i^{'} + j^{'} - 1 = n \implies \varepsilon \in R_n$. $\square$

THEOREM 4.8. (**Teorema di Sardinas e Patterson**). *Sia* $X \subseteq A^+$. X *non è codice* $\iff \exists n \geq 1 : X \cap R_n \neq \oslash$.

DIMOSTRAZIONE. La dimostrazione di questo teorema procedo attraverso i due lemmi precedenti. Prima si dimostrerà che X non è codice $\iff \varepsilon \in R_n$, grazie all'aiuto del lemma 4.7 a pagina 17e poi, grazie al lemma 4.6 a pagina 17 si arriverà a dimostrare la tesi di questo teorema.

Dimostriamo quindi che X non è codice $\iff \varepsilon \in R_n$.

($\implies$). Poiché X non è un codice $\exists x, y, x_1, ..., x_i, x_1^{'}, ..., x_j^{'}$ con $x \neq y$ tali che
$$x x_1 x_2 ... x_i = y x_1^{'} x_2^{'} ... x_j^{'}$$
Applicando il Lemma di Levi otteniamo due casi:
1) $| x | \geq | y |$
2) $| y | \geq | x |$
Caso 1: per il lemma di Levi otteniamo
$$\begin{cases} y = xp \\ x_1 ... x_i = p x_1^{'} ... x_j^{'} \end{cases} \implies \begin{cases} p \in R_1 \\ p X^j \cap X^i \neq \oslash \end{cases} \implies$$ (per il lemma 4.7 a pagina 17 con $k = 1$) $\varepsilon \in R_n$

Caso 2: per il lemma di Levi otteniamo
$$\begin{cases} x = yp \\ p x_1 ... x_i = x_1^{'} ... x_j^{'} \end{cases} \implies \begin{cases} p \in R_1 \\ X^j \cap p X^i \neq \oslash \end{cases} \implies$$ (per il lemma 4.7 a pagina 17 con $k = 1$) $\varepsilon \in R_n$

($\impliedby$). Dato che $\varepsilon \in R_n$, per il lemma 4.7 a pagina 17 con $k = 1$, $\exists r_1 \in R_1 \wedge \exists i, j$: $i + j + 1 = n \wedge r_1 X^i \cap X^j \neq \oslash$.

Poiché $r_1 \in R_1 = X^{-1} X \setminus \{\varepsilon\}$, $\exists x, y \in X : x r_1 = y$. Essendo $r_1 \neq \varepsilon$ per definizione di resto detro, $x \neq y$ in quanto $| x | + | r_1 | = | y |$, con $| r_1 | \geq 1$.

Dal fatto che $r_1 X^i \cap X^j \neq \oslash$ sappiamo che $r_1 x_1 ... x_i = x_1^{'} ... x_j^{'}$. Moltiplicando a sinistra ambo i membri di questa uguaglianza per x, otteniamo
$$x r_1 x_1 ... x_i = x x_1^{'} ... x_j^{'}$$
Abbiamo detto che $x r_1 = y$ quindi

$$yx_1...x_i = xx_1^{'}...x_j^{'}$$

Poiché $x_r \in X, \forall r = 1..i$ e $x_s^{'} \in X, \forall s = 1..j$ e $x, y \in X, x \neq y$ abbiamo trovato una parola di X che ha due fattorizzazioni diverse. Questo implica che X non è un codice.

Ora che si è dimostrato che X non è codice $\iff \varepsilon \in R_n$, per il lemma 4.6 a pagina 17, $\varepsilon \in R_n \iff R_{n-1} \cap X \neq \oslash$ e quindi si ottiene che $\exists n \geq 1 : R_n \cap X \neq \oslash \iff X$ non è un codice. □

Definition 4.9. (**Suffisso**). Sia X un insieme. Sia $x \in X$. Indicheremo con $Suff(x)$ un elemento s tale che $\exists y \in X : ys = x$.

Definition 4.10. (**Insieme dei suffissi**). Dato un insieme X indicheremo con $Suff(X)$ l'insieme dei suoi suffissi, così definito

$$Suff(X) = \{s \mid \exists x \in X : s = Suff(x)\}$$

Proposition 4.11. *Sia X un insieme e sia $x \in X$. Allora $\mid Suff(x) \mid \leq \mid x \mid +1$*[10].

Teorema di Levenstein.

Lemma 4.12. *Sia $X \subseteq A^+$ un insieme finito. Allora $\forall n \geq 1, R_n \subseteq Suff(X)$.*

Dimostrazione. La dimostrazione procede per induzione su n.

Passo base ($n = 1$). Consideriamo un elemento arbitrario di R_1; $w \in R_1 \iff \exists x_1, x_2 \in X$ con $x_1 \neq x_2 : x_1 w = x_2$, ma w è suffisso di x_2 quindi $w \in Suff(X)$. Dall'arbitrarietà di w segue che ciò vale per ogni elemento di R_1, quindi $R_1 \subseteq Suff(X)$.

Passo di induzione ($n > 1$). Si supponga vera l'ipotesi per n; dimostreremo che la tesi è vera anche per $n + 1$.

Consideriamo un aribitrario elemento $w \in R_{n+1}$; per definizione, $R_{n+1} = X^{-1}R_n \cup R_n^{-1}X \implies \exists x \in X \wedge \exists r_n \in R_n$ tali che

1) $xw = r_n$
2) $r_n w = x$

Nel secondo caso, w è suffisso di $x \in X$ il che implica che $w \in Suff(X)$. Nel caso 1 w è suffisso di $r_n \in R_n$ che per ipotesi induttiva è suffiso di una parola di X. Questo implica ancora una volta che $w \in Suff(X)$. Dall'arbitrarietà di w segue che ciò vale per ogni elemento di R_{n+1}, quindi $R_{n+1} \subseteq Suff(X)$. □

Lemma 4.13. *Sia $X \subseteq A^+$ un insieme finito. Sia $N = Card(X)$ e sia $L = \max_{x \in X} \{\mid x \mid\}$. Allora vale la seguete disuguaglianza*

$$Card(Suff(X)) \leq (NL) + 1$$

Dimostrazione. $Card(Suff(X)) = Card(\bigcup_{x \in X} Suff(x))$

Per la proposizione 4.11 sappiamo che $Suff(x) \leq \mid x \mid +1$.

Abbiamo dunque che $Card(\bigcup_{x \in X} Suff(x)) \leq \sum_{x \in X} \mid x \mid +1 \leq (NL) + 1$[11].

---

[10] Il "+1" dopo è per considerare la parola vuota ε.

[11] Il "+1" è fuori dalla sommatoria poiché la parola vuota va conteggiata una e una sola volta.

L'ultimo passaggio è possibile poiché la somma delle lunghezze di tutte le parole di un insieme è sicuramente maggiorabile con N volte la lunghezza della parola più lunga nell'insieme X di cardinalità N.

Pertanto, $Card(Suff(X)) \leq (NL) + 1$. $\square$

THEOREM 4.14. (**Teorema di Levenstein**). *Sia* $X \subseteq A^+$ *e sia* X *un insieme finito. Allora* X *è codice* $\iff \forall n, 1 \leq n \leq (NL) + 1\ R_n \cap X = \oslash$.

DIMOSTRAZIONE. ($\Longrightarrow$) X è codice per ipotesi. Per il teorema di Sardinas e Patterson[12] (4.8 a pagina 19) l'asserto segue banalmente.

($\Longleftarrow$) Supponiamo per assurdo che X non sia un codice. Allora per il teorema di Sardinas e Patterson $\exists n \geq 1 : R_n \cap X \neq \oslash$. Se $n \leq (NL) + 1$, allora si avrebbe una contraddizione, poiché per ipotesi $R_n \cap X = \oslash$ per $n \leq (NL) + 1$. Pertanto si suppone che $n > (NL)+1$ e tra tutti gli n che soddisfano tale condizione scegliamo il più piccolo, ovvero $\min\{n \mid n > (NL) + 1 \wedge R_n \cap X \neq \oslash\}$. Poiché si è supposto $R_n \cap X \neq \oslash$, allora R_n contiene almeno un elemento.

Sia $r_n \in R_n \implies \exists r_{n-1} \in R_{n-1}$ e $\exists x_{n-1} \in X$ tali che

$$\begin{cases} x_{n-1}r_n = r_{n-1} \\ r_{n-1}r_n = x_{n-1} \end{cases}$$

e quindi $\exists r_{n-1} \in R_{n-1}$ e $\exists x_{n-2} \in X$ e $\exists r_{n-2} \in R_{n-2}$ tali che

$$\begin{cases} x_{n-2}r_{n-1} = r_{n-2} \\ r_{n-2}r_{n-1} = x_{n-2} \end{cases}$$

e così via fino ad arrivare ai resti del primo ordine dove si avrà

$$\begin{cases} x_1r_2 = r_1 \\ r_1r_2 = x_1 \end{cases}$$

Si è così mostrato che per ogni $i \leq n$ esiste R_i non vuoto.

Avendo garantito l'esistenza degli insiemi R_i, si può creare una successione di resti destri $r_1r_2...r_n$ e di parole $x_1x_2...x_n$ tali che $r_i \in R_i$ e $x_i \in X$ per ogni $i = 1..n$ e

$$\begin{cases} x_ir_{i+1} = r_i \\ r_ir_{i+1} = x_i \end{cases}$$

Poiché ogni insieme R_i con $1 \leq i \leq n$ è non vuoto e $n > (NL) + 1$, allora $\bigcap_{i=1}^n R_i \neq \oslash$, poiché altrimenti sarebbe vero che $Card(\bigcup_{i=1}^n R_i) > NL + 1$, ma ciò è assurdo perché per il lemma 4.12 nella pagina precedente si ha che $\forall i, 1 \leq i \leq n\ R_i \subseteq Suff(X) \implies \bigcup_{i=1}^n R_i \subseteq Suff(X)$ e quindi $Card(\bigcup_{i=1}^n R_i) \leq Card(Suff(X)) \leq (NL) + 1$. Quest'ultimo passaggio è giustificato dal lemma 4.13 nella pagina precedente.

Assodato, dunque, che $\bigcap_{i=1}^n R_i \neq \oslash$, allora esistono $i, j \in \mathbb{N}$ tali che $1 \leq i < j \leq n$ e tali che $r_i = r_j$ con $r_i \in R_i$ e $r_j \in R_j$. Dalla successione costruita più sopra si ottiene

$$\begin{cases} x_jr_{j+1} = r_j \\ r_jr_{j+1} = x_j \end{cases}$$

essendo $r_j = r_i \in R_i$ allora la precedente può essere scritta come

$$\begin{cases} x_ir_{j+1} = r_i \\ r_ir_{j+1} = x_i \end{cases}$$

il che implica che $r_{j+1} \in R_{i+1}$. Reiterando il ragionamento si ha che

[12] Il teorema di Sardinas e Patterson afferma che X *non è codice se e soltanto se* $\exists n > 0 : R_n \cap X \neq \oslash$. Applicando l'operatore logico $\neg$ all'intera formulazione si ottiene ovviamente che X *è codice se e soltanto se* $\forall n > 0 R_n \cap X = \oslash$.

$$\Longrightarrow \begin{array}{c} r_{j+2} \in R_{i+2} \\ r_{j+3} \in R_{i+3} \\ \vdots \\ r_{j+n-j} \in R_{i+n-j} \end{array}$$

Sia $m = i + n - j = n - (j - i)$. Abbiamo dunque $r_{j+n-j} = r_n \in R_m$ e si è quindi trovato un elemento r_n tale che $r_n \in R_n \wedge r_n \in R_m$ e cioè un altro indice $m < n$ per il quale vale $R_m \cap X \neq \oslash$; ciò va a contraddire l'ipotesi di n come più piccolo indice per il quale $R_n \cap X \neq \oslash$. Si è giunti quindi a una contraddizione e pertanto segue che X è codice. $\square$

PROPOSITION 4.15. *Sia $X \subseteq A^+$ e sia X un insieme finito; allora è possibile decidere se X è un codice o meno.*

DIMOSTRAZIONE. Sia $R_1, R_2, ..., R_n, ...$ la successione dei resti destri di X. Per il lemma 4.12 a pagina 20 ogni resto destro è contenuto in $Suff(X)$ che è un insieme finito, quindi $\exists i, j \in \mathbb{N} : 0 < i < j$ per i quali $R_i = R_j$. Supponiamo che j sia il più piccolo valore per il quale $\exists i : (i < j \wedge R_i = R_j)$. A questo punto si costruisce la successione $R_1, R_2, ..., R_{j-1}$. Se $R_j = R_i$ allora $R_{j+1} = R_{i+1}$ poiché $R_{j+1} = X^{-1}R_j \cup R_j^{-1}X = X^{-1}R_i \cup R_i^{-1}X = R_{i+1}$.

Questo significa che si ha un insieme di resti destri finito. $\square$

DEFINITION 4.16. Dati $X, Y \subseteq A^+$, $X \cdot Y = \{xy \in A^* \mid x \in X \wedge y \in Y\}$. Il prodotto $X \cdot Y$ è detto non ambiguo se $\forall w \in X \cdot Y\ \exists!(x, y) \in X \cdot Y : w = xy$.

PROPOSITION 4.17. *Sia $X \subseteq A^*$. X è codice $\iff$ $\forall n > 0\ X \cdot X^n$ è non ambiguo.*

DIMOSTRAZIONE. ($\Longrightarrow$) Supponiamo per assurdo che $\exists n : XX^n$ è ambiguo. Allora, per definzione di *ambiguo* $\exists x, y \in X$ e $\exists w_1, w_2 \in X^n$ tali che $xw_1 = yw_2$ con $w_1 = x_1...x_n$ e $w_2 = y_1...y_n$, quindi $xx_1...x_n = yy_1...y_n$, il che implica che esiste una parola che ha due fattorizzazioni e ciò è assurdo poiché per ipotesi X è un codice.

($\Longleftarrow$) Supponiamo per assurdo che X non sia codice, allora $\exists x, y \in X$ e $\exists x_1, ..., x_h, x_1^{'}, ..., x_k^{'}$ tali che $xx_1...x_h = yx_1^{'}...x_k^{'}$ il che ci permette di scrivere

$$xx_1...x_h y x_1^{'}...x_k^{'} = yx_1^{'}...x_k^{'} x x_1...x_h$$

con $x \in X$, $x_1...x_h y x_1^{'}...x_k^{'} \in X^{h+1+k}$, $y \in X$ e $x_1^{'}...x_k^{'} x x_1...x_h \in X^{k+1+h}$, quindi questo porterebbe all'ambiguità di XX^{h+k+1}, che assurdo per le ipotesi di non ambiguità. $\square$

PROPOSITION 4.18. *Sia $X \subseteq A^+$ allora X è prefisso $\iff$ XA^* è non ambiguo.*

Distribuzioni di Bernoulli.

DEFINITION 4.19. Una **distribuzione di Bernoulli** è una funzione ρ da un insieme A all'insieme $\mathbb{R}^+$, $\rho : A \to \mathbb{R}^+$, tale che $\rho(a) \geq 0\ \forall a \in A$ e tale che $\sum_{a \in A} \rho(a) = 1$. Tale distribuzione è detta positiva se $\forall a \in A, \rho(a) > 0$.

Diseguaglianza di Kraft-McMillan. La diseguaglianza di Kraft-McMillan è un caso particolare del teorema di Kraft-McMillan per distribuzioni uniformi.

Sia X un codice su un alfabeto A di cardinalità d, allora si ha che

$$\sum_{x \in X} d^{-|x|} \leq 1$$

La diseguaglianza afferma, in pratica, che se X è codice non è possibile avere molte parole di lunghezza piccola.

Definition 4.20. (**Funzione di struttura**). Sia X un codice su A. Si definisce funzione di distribuzione della lunghezza o funzione di struttura di X l'applicazione $f_x : \mathbb{N} \to \mathbb{N}$ tale che $f_X(n) = Card(X \cap A^n)$.

La funzione di struttura permette di sapere quante parole di lunghezza n sono presenti nell'insieme X. La diseguaglianza di Kraft-McMillan in termini della funzione di distribuzione della lunghezza si presenta come segue

$$\sum_{n \geq 0} f_x(n) d^{-n} \leq 1$$

Definition 4.21. Una distribuzione di Bernoulli **uniforme** su un insieme A con $Card(A) = d$ ha come caratteristica che $\forall a \in A, \rho(a) = \frac{1}{d}$.

Definition 4.22. (**Estensione della Distribuzione di Bernoulli dalle lettere alle parole**). Sia $\rho : A \to \mathbb{R}^+$ una distribuzione di Bernoulli. Essa si estende univocamente a un morfismo $\hat{\rho} : A^* \to \mathbb{R}^+$ con $\hat{\rho}(\varepsilon) = 1$ e $\forall u, v \in A^*, \hat{\rho}(uv) = \hat{\rho}(u)\hat{\rho}(v)$.

Definition 4.23. (**Estensione della Distribuzione di Bernoulli alle parti**). Sia $\mathcal{B}(A^*)$ l'insieme delle parti di A^*, ovvero $\mathcal{B}(A^*)$ è l'insieme di tutti i linguaggi sull'alfabeto A.

$\rho : \mathcal{B}(A^*) \to \mathbb{R}^+ \cup \{+\infty\}$ con ρ definita come segue:
1. $\rho(\oslash) = 0$
2. $\rho(X) = \sum_{x \in X} \rho(x)$

Se X è finito, allora $\rho(X) < +\infty$, se invece X è un insieme infinto, allora $\rho(X)$ può essere finito o infinito.

In aggiunta, ρ verifica la seguente proprietà

$$\rho(\bigcup_i X_i) \leq \sum_i \rho(X_i)$$

L'uguaglianza si ottiene nel caso in cui $\bigcap_i X_i = \oslash$.

Definition 4.24. Definiamo DB come l'insieme delle Distribuzioni di Bernoulli e definiamo $DBP \subset DB$ l'insieme delle Distribuzioni di Bernoulli positive.

Proposition 4.25. *Siano $X, Y \in \mathcal{B}(A^*)$ e $\rho \in DB$, allora si ha che $\rho(xy) \leq \rho(x)\rho(y)$. Se $X \cdot Y$ è non ambiguo, allora $\rho(XY) = \rho(X)\rho(Y)$.*

Dimostrazione. $X = \bigcup_{x \in X} \{x\}$ e $Y = \bigcup_{y \in Y} \{y\}$ allora $X \cdot Y = \bigcup_{x \in X} \{x\} \cdot \bigcup_{y \in Y} \{y\} = \bigcup_{x \in X, y \in Y} \{x \cdot y\}$.

Questo implica che $\rho(X \cdot Y) = \rho(\bigcup_{x \in X, y \in Y} \{xy\}) \leq \sum_{x \in X, y \in Y} \rho(\{xy\}) = \sum_{x \in X, y \in Y} \rho(x)\rho(y) = \sum_{x \in X} \rho(x) \cdot \sum_{y \in Y} \rho(y) = \rho(x)\rho(y)$. $\square$

Proposition 4.26. *Se $\rho \in DBP$ e $\rho(XY) = \rho(X)\rho(Y)$ e ancora $\rho(XY) < +\infty$ allora XY è non ambiguo.*

Dimostrazione. Supponiamo per assurdo che $X \cdot Y$ sia ambiguo, allora $\exists w : w = xy = x^{'}y^{'}$ e $\rho(w) > 0$. Stando così le cose, allora $\rho(X)\rho(Y) \geq \rho(XY) + \rho(w)$; poiché $\rho(XY) < +\infty$, si ottiene che $\rho(X)\rho(Y) > \rho(XY)$ il che ci conduce a un assurdo. Pertanto XY non è ambiguo. $\square$

Proposition 4.27. *Sia $X \subseteq A^+$ e sia $\rho \in DB$. Se X è codice, allora $\forall n > 0$ $\rho(X^n) = [\rho(X)]^n$.*

Dimostrazione. Per ipotesi X è un codice, allora per la proposizione 4.18 XX^n non è ambiguo e si può quindi scrivere $\rho(X)\rho(X^n) = \rho(XX^n)$ essendo $\rho \in DB$.

Procediamo ora per induzione su n.

Passo base ($n = 1$). $n = 1 \implies \rho(X) = [\rho(X)]^1$ **-** banale.

Passo di induzione. Supponiamo l'ipotesi vera per n e dimostriamo la stessa per $n + 1$.

$\rho(X^{n+1}) = \rho(XX^n) = \rho(X)\rho(X^n) =_{per\, ipotesi\, induttiva} \rho(X)[\rho(X)]^n = [\rho(X)]^{n+1}$ - c.v.d. $\square$

Proposition 4.28. *Sia $X \subseteq A^+$ e sia $\rho \in DB$. Se $\rho \in DBP$, $\rho(X^n) = [\rho(X)]^n$, $\rho(X) < +\infty$, allora X è un codice.*

Dimostrazione. Per dimostrare che X è un codice, bisogna mostrare che $\forall n$ XX^n è non ambiguo. Ricordando che $\rho(XY) = \rho(X)\rho(Y) \iff$ *XY è non ambiguo*, possiamo mostrare che $\rho(XX^n) = \rho(X)\rho(X^n)$:

$\rho(XX^n) = \rho(X^{n+1}) =_{per\, ipotesi} [\rho(X)]^{n+1} = \rho(X)[\rho(X)]^n = \rho(X)\rho(X^n)$.

Ora ci resta da mostrare che $\rho(XX^n) < +\infty$:

$\rho(XX^n) = \rho(X^{n+1}) = [\rho(X)]^{n+1} < \infty$, essendo $\rho(X) < +\infty$ per ipotesi.

Segue che X è codice. $\square$

Proposition 4.29. *Sia A un alfabeto. Allora $\forall n \geq 0$ $\forall \rho \in DB$, $\rho(A^n) = 1$.*

Dimostrazione. $\rho(A^n) = [\rho(A)]^n$ ma essendo $\rho(A) = 1$ in quanto $\rho \in DB$, allora si ha $[\rho(A)]^n = [1]^n = 1$ $\square$

Theorem 4.30. (**Teorema di Kraft-McMillan generalizzato**). Sia X un codice su A. Allora $\forall \rho \in DB$, $\rho(X) \leq 1$.

Dimostrazione. (*Caso di X finito*). Se X è finito, è possibile prendere

$$L = \max\{\mid x \mid : x \in X\}$$

e inoltre sappiamo che

$$X \subseteq A \cup A^2 \cup ... \cup A^L = \bigcup_{i=1}^{L} A^i$$

La lunghezza delle parole su X è limitata da $n \cdot L$ e cioè

$$X^n \subseteq \bigcup_{i=1}^{nL} A^i$$

ovvero la parola più lunga che è possibile ottenere è costituita da n volte la concatenazione della parola di lunghezza L, quindi

$$\rho(X^n) \leq \rho(\bigcup_{i=1}^{nL} A^i) \leq \sum_{i=1}^{nL} \rho(A^i)$$

Per la proposizione 4.29 la precedente è uguale a

$$\sum_{i=1}^{nL} 1 = nL$$

ovvero $\rho(X^n) \leq nL \implies [\rho(X)]^n \leq nL \implies \rho(X) \leq n^{1/n}L^{1/n}$, quindi $\rho(X) \leq \lim_{n\to+\infty} n^{1/n}L^{1/n} = 1$.

(*Caso di X infinito*). Consideriamo $X_k = \{x \in X :\mid x \mid\leq k\}$; ogni $X_k \subseteq X$ è codice[13] finito, il che implica, da quanto dimostrato sinora, che $\rho(X_k) \leq 1$. Poiché $X_k \subseteq X_{k+1}$ si ha che $\rho(X_k) \leq \rho(X_{k+1}) \leq 1$. Si crea così una successione monotona crescente a termini non negativi e pertanto vale

$$\lim_{k\to+\infty} \rho(X_k) \leq 1$$

Sapendo che

$$X = X_1 \cup (X_2 \backslash X_1) \cup (X_3 \backslash X_2) \cup ... \cup X_{i+1} \backslash X_i) \cup ... = X_1 \cup \bigcup_{i=1}^{n} (X_{i+1} \backslash X_i)$$

passando alle probabilità si ottiene

$$\rho(X) = \rho(X_1) + \sum_{i=1}^{+\infty} \rho(X_{i+1} \backslash X_i)$$

ma $\rho(X_{i+1} \backslash X_i) = \rho(X_{i+1}) - \rho(X_i)$, perciò sostituendo nella precedente

$$\rho(X) = \rho(X_1) + \sum_{i=1}^{+\infty} \rho(X_{i+1}) - \rho(X_i) = \rho(X_1) + \lim_{n\to+\infty} \sum_{i=1}^{n} \rho(X_{i+1}) - \rho(X_i)$$

Sviluppando la sommatoria otteniamo

$$\rho(X_{n+1}) - \rho(X_n) + \rho(X_n) - \rho(X_{n-1}) + ... + \rho(X_2) - \rho(X_1) = \rho(X_{n+1}) - \rho(X_1)$$

quindi

$$\rho(X) = \rho(X_1) + \lim_{n\to+\infty} [\rho(X_{n+1}) - \rho(X_1)] = \rho(X_1) - \rho(X_1) + \lim_{n\to+\infty} \rho(X_{n+1}) = \lim_{n\to+\infty} \rho(X_{n+1}) \leq 1$$

□

[13] Si osservi che se X è un codice e $Y \subseteq X$ allora Y è anch'esso un codice.

Completezza e massimalità.

DEFINITION 4.31. Sia X un codice su A. X è massimale se per ogni altro Y codice su A si ha che $X \subseteq Y \implies X = Y$.

PROPOSITION 4.32. *Sia X un codice su A e sia $\rho \in DBP$. Se $\rho(X) = 1$ allora X è massimale.*

DIMOSTRAZIONE. Per ipotesi $\rho(X) = 1$. Sia Y un altro codice su A e si supponga $X \subset Y$, allora $\rho(Y) = \rho(X) + \rho(Y \backslash X) = 1 + \rho(Y \backslash X)$. Essendo $\rho \in DBP$, $\rho(Y \backslash X) > 0$, quindi $\rho(Y \backslash X) + 1 > 1$, che è assurdo, quindi $X = Y$, ovvero X è massimale. $\square$

DEFINITION 4.33. Sia $X \subseteq A^*$. X è denso $\iff \forall f \in A^* \exists u, v \in A^* : ufv \in X$, cioè $A^* f A^* \cap X \neq \oslash$.

PROPOSITION 4.34. *Se X è denso e $\rho \in DBP$ allora $\rho(X) < +\infty$.*

PROPOSITION 4.35. *Se X è un codice regolare, allora X è non denso.*

PROPOSITION 4.36. *Sia $X \subseteq A^*$. Se X^* è denso, allora X è completo.*

LEMMA 4.37. (**Lemma di Lyndon-Schützenberger**). *Siano $f, g \in A^+$ e $h \in A^*$ tali che $fh = hg$; allora $\exists \lambda, \mu \in A^*$ e $\exists n \in \mathbb{N}$ tali che $f = \lambda\mu$ e $g = \mu\lambda$ e $h = (\lambda\mu)^n \lambda$.*

DIMOSTRAZIONE. La dimostrazione procede per induzione su $| h |$.
Passo base. $| h |= 0 \implies h = \varepsilon$ e quindi $f = g$; allora poniamo $\lambda = \varepsilon$, $n = 0$ e $\mu = f = g$.
Passo di induzione.
Si supponga vera la tesi per un h di una determinata lunghezza.
Per ipotesi abbiamo $fh = hg$ e possono verificarsi, come sempre, due casi
1) $| f |\geq| h |$
2) $| h |>| f |$
Caso 1. Per il Lemma di Levi (3.14 a pagina 11): $\exists p \in A^* : f = hp \implies fh = hph \implies hg = hph \implies g = ph$.
Dato che $ph = g$, basta scegliere $\lambda = h$, $\mu = p$, $n = 0$.
Caso 2. Per il lemma di Levi $\exists z \in A^* : h = fz$. Sia $z = h^{'}$ con $| h' |<| h |$. Questo implica che $ffh^{'} = fh^{'}g \implies fh^{'} = h^{'}g$. Essendo $| h^{'} |<| h |$, è possibile applicare l'ipotesi induttiva e pertanto $\exists \mu, \lambda$ e $\exists n \in \mathbb{N} : f = \lambda\mu$, $g = \mu\lambda$, $h^{'} = (\lambda\mu)^n \lambda$. A questo punto non resta che determinare la h: $h = fh^{'} = (\lambda\mu)(\lambda u)^n \lambda = (\lambda\mu)^{n+1} \lambda$. $\square$

PROPOSITION 4.38. (**Proposizione di Marcus-Schützenberger**). Sia X un insieme non denso e completo, allora $\forall \rho \in DBP$ $\rho(X) \geq 1$.

DIMOSTRAZIONE. Essendo X non denso, per definizione esiste almeno una parola che non si completa in X ovvero $\exists w \in A^* : A^* w A^* \cap X = \oslash$. Essendo X completo, allora X^* è denso e quindi ogni parola di X^*si completa in A^*, quindi $\forall f \in X^*$ si può considerare la parola wfw che si completa in X^* ovvero $A^* wfw A^* \cap X^* \neq \oslash$; ciò implica che $\exists u, v \in A^* : uwfwv \in X^*$ o ancora che

$\exists n > 0 \exists x_1 ... x_n \in X^* \exists u, v \in A^*$ tali che $uwfwv = x_1...x_n$. Chiaramente non può esserci un elemento x tale che $w \leq x$ perché altrimenti w si completerebbe in X. Questo implica che per w passa almeno una linea di parsing.

Ciò implica che $w = w_1 w_2 = w_3 w_4$, ovvero w viene suddivisa in due dalla linea di parsing.

Poiché la linea di parsing inizia su w_2 e finisce su w_3, la parola $w_2 f w_3$ può essere scritta come prodotto di elementi di X, ovvero: $w_2 f w_3 \in X^*$ e inoltre $w_1 \in PREF(w) = P_w$ e $w_4 \in SUFF(w) = S_w$.

Abbiamo dunque $wfw = w_1 w_2 f w_3 w_4 \in w_1 X^* w_4 \subseteq P_w X^* S_w$ ed essendo f arbitraria si può scrivere $w_1 X^* w_4 \subseteq P_w X^* S_w$ come $wA^*w \subseteq P_w X^* S_w$.

Sia $\rho \in DBP$, allora $\rho(w) > 0$ e inoltre $\rho(A^*) = +\infty$ quindi ricordando che $wA^*w \subseteq P_w X^* S_w$, possiamo scrivere

$$+\infty = \rho(w_1 A^* w) = \rho(w^2)\rho(A^*) \leq \rho(P_w X^* S_w) = \rho(P_w)\rho(X^*)\rho(S_w)$$

Essendo $\rho \in DBP$ ed essendo P_w e S_w degli insieme finiti, si hanno due condizioni importanti:

1) $\rho(P_w) > 0$ e $\rho(S_w) > 0$

2) $\rho(P_w)\rho(S_w) < +\infty$

Di conseguenza la relazione precedente è vera se e solo se $\rho(X^*) = +\infty$, quindi

$$+\infty = \rho(X^*) = \rho(\bigcup_{n=0}^{+\infty} X^n) \leq \sum_{n=0}^{+\infty} \rho(X^n) = \sum_{n=0}^{+\infty} \rho(X)^n$$

Affinché la serie diverga deve valere $\rho(X) \geq 1$. - c.v.d - $\square$

THEOREM 4.39. (**Secondo Teorema di Schützenberger**). Sia X un codice su A.

1) Se X è un codice massimale, allora X è completo.

2) Se X è un codice completo e non denso, allora X è massimale.

DIMOSTRAZIONE. (1). Supponiamo per assurdo che X non sia completo, allora $\exists f \in A^* : A^* f A^* \cap X^* = \oslash$. Si facciano due ulteriori ipotesi:

a) $Card(A) > 1$

b) f è una parola primaria, cioè $f \neq uvu$ con $u \neq \varepsilon$ e $u, v \in A^*$

Sia $Y = X \cup \{f\}$ dove f, ribadiamo, è la parola che non si completa in X^*.

La dimostrazione prosegue in questo modo: si cerca di dimostrare che Y è un codice ottenendo ogni volta un assurdo contraddicendo l'ipotesi di f come parola che non si completa in X^*. Ciò significa che non esiste alcuna parola che non si completa in X^*, quindi X è completo.

Vogliamo dunque dimostrare che Y è un codice e supponiamo per assurdo che Y non lo sia affatto; allora $\exists y_1, ..., y_h, y_1^{'}, ..., y_k^{'}$ tali che $y_1 \neq y_1^{'}$, $y_1...y_h = y_1^{'}...y_k^{'}$ e $y_i, y_j^{'} \in Y, \forall i = 1..h, \forall j = 1..k$.

L'unico elemento in Y che non è in X è la parola f. Dunque si possono verificare tre casi:

1) f non si trova né a sinistra né a destra dell'uguaglianza $y_1...y_h = y_1^{'}...y_k^{'}$

2) f occorre solo a sinistra o solo a destra dell'uguaglianza;

3) f occorre sia a sinistra sia a destra dell'uguaglianza, anche più volte.

Caso 1. Questa situazione implicherebbe che X non è un codice, che è una contraddizione;

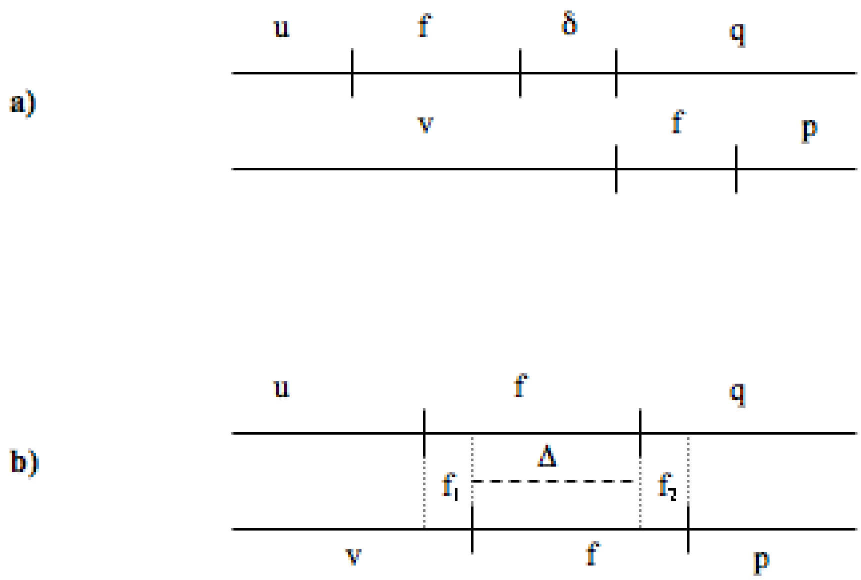

FIGURA 1. a) Sottocaso 1; b) Sottocaso 2.

Caso 2. Se f occorre solo a sinistra, nella parte destra della uguaglianza ci sono soltanto elementi di X; questo implica che $\forall i \ y_i^{'} \in X$ e quindi si ha

$$y_1..y_{i-1}fy_{i+1}...y_h = y_1^{'}...y_k^{'}$$

con $y_1..y_{i-1} = u \in A^*$, $y_{i+1}...y_h = v \in A^*$ e $y_1^{'}...y_k^{'} = x \in X^*$, abbiamo quindi $ufv \in X^*$. Questa è una contraddizione in quanto per ipotesi f è la parola che non si completa in X^*.

In modo analogo si mostra il sottocaso di f che occorre solo a destra.

Caso 3. La parola f occorre sia a sinistra sia a destra. Isoliamo le prime occorrenze di f in modo che tutte le lettere che precedono f appartengono a X^*

$$y_1..y_{i-1}fy_{i+1}...y_h = y_1^{'}...x_{j-1}^{'}fy_{j+1}^{'}...y_k^{'}$$

con $y_1..y_{i-1} = u \in X^*$, $y_{i+1}...y_h = q \in Y^*$, $y_1^{'}...y_{j-1}^{'} = v \in X^*$ e $y_{j+1}^{'}...y_k^{'} = p \in Y^*$, quindi

$$ufq = vfp$$

A questo punto possiamo avere due sottocasi.

Sottocaso 1. $\mid v \mid \geq \mid uf \mid$ (o analogamente $\mid u \mid \geq \mid vf \mid$)

Per il Lemma di Levi segue che $\exists \delta : uf\delta = v$ che implica che $A^*fA^* \cap X^* \neq \oslash$; ciò è assurdo in quanto per ipotesi f non si completa.

Sottocaso 2. La f si accavalla, ovvero abbiamo $f = f_1\Delta = \Delta f_2$ da cui, per il lemma di Lyndon-Schützenberger (4.37 a pagina 26) $\exists \lambda, \mu \in A^*$ e $\exists n \in \mathbb{N}$ tali che

$f_1 = \lambda\mu$, $f_2 = \mu\lambda$ e $\Delta = f_3 = (\lambda\mu)^n\lambda$. Questo implica che $f = f_1\Delta = \lambda\mu(\lambda\mu)^n\lambda$, ma f è primaria, quindi deve essere $\lambda = \varepsilon$. Allora $f_1 = \mu = f_2$ e $\Delta = (\mu)^n$; ciò implica che $f = \mu(\mu)^n$ ma essendo f primaria, abbiamo $n = 0$ in quanto f non può iniziare e finire con $\mu \neq \varepsilon$. Segue che $\Delta = \varepsilon$, ovvero che non c'è alcun accavallamento tra le f e quindi il sottocaso in questione non può verificarsi.

A questo punto cambiamo l'ipotesi b) e supponiamo che f non sia primaria. Allora possiamo costruire $g = fb^{|f|}$ tale che $b \in A$ sia diversa dalla prima lettera di f[14]. La parola g è quindi costruita in modo che:

i)se f non si completa in X^* allora g non si completa in X^* .

ii)g è primaria

Si usa la dimostrazione precedente con il codice $X \cup \{g\}$.

Dimostraiamo la i): se g si completasse in X^* allora $\exists u, v \in A^*$ tale che $ugv \in X^* \implies ufb^{|f|}v \in X^* \implies f$ si completa in X^*.

Dimostraimo la ii): g è primaria perché se non lo fosse $\exists u, v \in A^*$ tali che $g = uvu$ ma per costruzione ciò non è possibile, poiché la prima u inizia con una lettera diversa da b e la seconda u, per costruzione, può iniziare a partire da destra con una lettera diversa da b solo dopo $| f |$ lettere. Ciò significherebbe che $| u_{dx} |>| f | \Longrightarrow | u_{sx} | + | u | + | u_{dx} |> 2 | f |>| g | \implies u_{sx}vu_{dx} \neq g$.

(2). Per il teorema di Kraft-McMillan (4.30 a pagina 24) abbiamo che X codice $\implies \rho(X) \leq 1$. Per la proposizione di Marcus-Schützenberger (4.38 a pagina 26) abbiamo che X non denso $\implies \rho(X) \geq 1$.

Quindi $\rho(X) = 1$ e per la proposizione 4.32 a pagina 26, X è codice massimale.

□

Definition 4.40. Siano $f, g \in A^*$, f si dice coniugata g ($f\mathcal{C}g$) se $\exists u, v \in A^*$: $f = uv \wedge g = vu$.

La coniugazione è una relazione di equivalenza su A^*.

Proposition 4.41. *Sia X un codice non denso. Le seguenti affermazioni sono equivalenti:*

1) *X è completo*

2) $\forall\mu \in DBP$, $\mu(X) = 1$

3) $\exists\mu \in DBP : \mu(X) = 1$

4) *X è massimale*

Dimostrazione. Per dimostrare la proposizione basta dimostrare che 1) $\implies$ 2) $\implies$ 3) $\implies$ 4) $\implies$ 1).

1) $\implies$ 2). X non denso e completo. Per la proposizione 4.38 a pagina 26 questo implica che $\forall\mu \in DBP, \mu(X) \geq 1$. Il fatto che X è codice, invece, implica (per il teorema di Kraft-McMillan), che $\forall\mu \in DBP, \mu(X) \leq 1$.

Per tanto, $\forall\mu \in DBP, \mu(X) = 1$.

2) $\implies$ 3). Banale.

3) $\implies$ 4). Per la proposizione 4.32 a pagina 26.

4) $\implies$ 1). Il fatto che X sia un codice non denso e massimale implica, per il teorema di Scützenberger (4.39 a pagina 27), che X è completo. □

---

[14]Esempio: sia $f = wxyz$ e sia $b \neq w$, allora $g = wxyzbbbb$.

PROPOSITION 4.42. *Sia $\mathcal{F}$ la famiglia di tutti i codici su A, $\mathcal{F} = \{X_\gamma\}_{\gamma\in\Gamma}$, con Γ insieme di indici. Allora $\mathcal{F}$ possiede la proprietà di Zorn: $\forall\gamma \in \Gamma, X_\gamma \subseteq X_{\gamma+1} \implies Y = \bigcup_{\gamma\in\Gamma} X_\gamma$ è un codice.*

DIMOSTRAZIONE. Dobbiamo dimostrare che Y è un codice, ovvero che

$$\forall y_i, y_i^{'} \in Y : y_1...y_h = y_1^{'}...y_k^{'} \implies \begin{cases} h = k \\ \forall i = 1..h, y_i = y_i^{'} \end{cases}$$

Ogni y apparterrà a un elemento della famiglia, ovvero $\forall y_i, y_i^{'} \exists X_\gamma : y_i \in X_\gamma \vee y_i^{'} \in X_\gamma$. Consideriamo $X_\gamma = \max\left\{X_\gamma : \exists i : y_i \in X_\gamma \vee y_i^{'} \in X_\gamma\right\}$, allora X_γ contiene tutte gli y_i e gli $y_i^{'}$ poiché X_γ è il massimom rispetto alla catena di inclusione. Ma X_γ è un codice (soddisfa la proprietà di fattorizzazione unica), dunque $\forall i, y_i = y_i^{'}$ e $h = k$ - c.v.d. □

PROPOSITION 4.43. *Sia X un codice su A. Allora X ha un completamento se esiste Y codice massimale su A tale che $X \subseteq Y$.*

COROLLARY 4.44. *Sia X un codice su A. Se X è massimale, allora X è completamento di se stesso.*

PROBLEM 4.45. Sia X un codice. Esiste sempre un completamenteo di X?

PROBLEM 4.46. Sias X un codice finito. Esiste sempre un completamento finito di X?

PROBLEM 4.47. Esiste un algoritmo in grando di decidere se un codice finito ha un completamento finito?

DEFINITION 4.48. (Ordine lessicografico).

$u \leq_{lex} v$ se $v \in uA^*$ oppure$\begin{cases} u = hX\xi \\ v = hY\eta \end{cases}$ con $h \in A^*$, $x, y \in A$, $\xi, \eta \in A^*$ e $x < y$

EXAMPLE. $fatto \leq_{lex} fattore \leq_{lex} fattorino$

DEFINITION 4.49. (Ordine militare).

$u \leq_{mil} v$ se $\mid u \mid < \mid v \mid$ oppure $\mid u \mid = \mid v \mid$ e $u \leq_{lex} v$

NOTE 4.50. La relazione di ordinamento è di buon ordine: ogni sottoinsieme ammette minimo.

PROPOSITION 4.51. *Sia A un alfabeto finito. Ogni codice su A ammette completamento.*

DIMOSTRAZIONE. Sia X un codice su A. Se X è massimale allora X è il completamento di se stesso.

Sia X, invece, non massimale. Allora $\exists w \in A^* : X \cup \{w\}$ è codice. Sia w_1 la più piccola parola w in ordine militare tale che $X \cup \{w_1\}$ è codice, allora $Y_1 = X \cup \{w_1\}$ è massimale oppure $\exists w : Y_1 \ \{w\}$ è codice. Sia w_2 la più piccola parola w in ordine militare tale che $Y_1 \cup \{w_2\}$ è codice, allora $Y_2 = Y_1 \cup \{w_2\}$ è massimale...

Continuando il ragionamento si ha $X = Y_0 \subseteq Y_1 \subseteq Y_2 ... \subseteq Y_n \subseteq ...$ con $Y_n = Y_{n-1} \cup \{w_n\}$.

Sia $Z = \bigcup_{n \geq 0} Y_n$, per la proprietà di Zorn (4.42 a pagina 29) Z è codice. Resta da dimostrare che Z è massimale.

Supponiamo per assurdo che Z non è massimale; allora $\exists \overline{w}$ tale che $Z \cup \{\overline{w}\}$ è codice. Ora, per come vengono scelte le parole w_i accade che:

$$w_1 \leq_{mil} w_2 \leq_{mil} ... \leq_{mil} w_n \leq_{mil} ...$$

Dalla definizione di ordine militare e dalla finitezza dell'alfabeto A (che implica che non ci sono infinite parole di lunghezza uguale) esisterà una sottosuccessione di w_{j_i} tale che

$$| w_{j_1} |<| w_{j_2} |< ... <| w_{j_r} |< ...$$

ovvero le parole w_{j_i} avranno lunghezza arbitraria.

A questo punto, essendo $\overline{w}$ una parola finita, esisterà un k tale che $| w_{k-1} |\leq| \overline{w} |<| w_k |$

Ma se $Z \cup \{\overline{w}\}$ è un codice, allora, essendo Y_{k-1} contenuto in Z, $Y_{k-1} \cup \{\overline{w}\}$ è un codice.

Ma $\overline{w}$ ha lunghezza minore di w_k; questo porta a un assurdo perché w_k è la più piccola parola secondo l'ordine militare tale che $Y_{k-1} \cup \{w_k\}$ è un codice. Quindi Z è massimale - c.v.d. □

5. Teoria dell'informazione

Definizione della sorgente in modo statistico.

DEFINITION 5.1. Sia $p \in DB$ e Y un alfabeto. Una sorgente è definita come la coppia $S = [Y, p]$ nella quale $p(y)$, con $y \in Y$, rappresenta la probabilità che la sorgente emetta la lettera y.

DEFINITION 5.2. Sia S una sorgente, sia X un codice su A e $\varphi : Y^* \to A^*$ un monomorfismo di codifica, il costo del morfismo è definito come

$$C(X, \varphi) = \sum_{y \in Y} p(y) \mid \varphi(y) \mid = \sum_{x \in X} p(x) \mid x \mid$$

con $p(x) = p(\varphi^{-1}(y))$.

Chiaramente, un morfismo ha costo minore quando associa codifiche più brevi a messaggi sorgente più frequenti.

DEFINITION 5.3. Sia $S = [Y, p]$ una sorgente, diremo che il codice X è adattato a S se esiste una biezione $\varphi : Y \leftrightarrow X$.

DEFINITION 5.4. Sia X un codice su A e $S = [Y, p]$ una sorgente. Definiamo costo assoluto di X

$$C(X) = \min \{C(X, \varphi)\}$$

con φ una qualsiasi biezione[15] $Y \leftrightarrow X$.

Con questa definizione il costo non dipende più dal morfismo φ.

PROPOSITION 5.5. *Sia $S = [Y, p]$ una sorgente e sia $C(X) = \min\{C(X, \varphi) : \varphi : Y \leftrightarrow X\}$. Se $C(X) = C(X, \varphi_0)$ allora $\forall x_i, x_j \in X\ p_i > p_j \implies \mid x_i \mid \leq \mid x_j \mid$.*

DIMOSTRAZIONE. Supponiamo per assurdo che $\exists x_i, x_j \in X : p_i > p_j$ e $\mid x_i \mid > \mid x_j \mid$. Allora possiamo costruire un nuovo morfisimo $\hat{\varphi}$ tale che

$$\hat{\varphi}(y) = \begin{cases} \varphi_0(y) & se\, y \neq y_i, y_j \\ x_j & se\, y = y_i \\ x_i & se\, y = y_j \end{cases}$$

con y_i e y_j le controimmagini, rispettivamente, di x_i e x_j.
Effettuiamo la differenza tra il costo di φ_0 e quello di $\hat{\varphi}$

$$\begin{aligned} C(X, \varphi_0) - C(X, \hat{\varphi}) &= \\ &= \sum_{l=1}^{n} p_l \mid x_l \mid -(\sum_{l=1, l \neq x_i, x_j} p_l \mid x_l \mid + p_i \mid x_j \mid + p_j \mid x_i \mid) = \\ &= p_i \mid x_i \mid + p_j \mid x_j \mid - p_i \mid x_j \mid - p_j \mid x_i \mid = \\ &= p_i(\mid x_i \mid - \mid x_j \mid) + p_j(\mid x_j \mid - \mid x_i \mid) = \\ &= p_i(\mid x_i \mid - \mid x_j \mid) - p_j(\mid x_i \mid - \mid x_j \mid) = \\ &= (p_i - p_j)(\mid x_i \mid - \mid x_j \mid) > 0 \end{aligned}$$

ovvero il costo $\hat{\varphi}$ è minore di quello di φ_0 , il che ci porta a un assurdo. □

DEFINITION 5.6. Sia $S = [Y, p]$ una sorgente, il codice X adattato a S si dice ottimale per S se qualsiasi altro codice Z adattato a S ha costo maggiore, ovvero

$$\forall Z, C(Z) \geq C(X)$$

Nel seguito dimostreremo che:

(1) esiste sempre un codice prefisso ottimale
(2) se $Card(Y) = 2$ e X è ottimale per S allora X è massimale

Entropia di una sorgente. Shannon definisce come la misura della quantità di informazione media associata a un risultato casuale. La base del logaritmo originariamente utilizzata da Shannon fu quella naturale, tuttavia è oggi di uso comune la base 2 in quanto consente di ottenere dei risultati più chiari (in particolare, il valore di entropia ottenuta è misurato in bit).
Nel seguito sarà utilizzata la base classica scelta da Shannon.

DEFINITION 5.7. Ad ogni sorgente S è associata una quantità $H(S)$ detta entropia della sorgente S

$$H(S) = \sum_{y \in Y} p(y) \ln \frac{1}{p(y)} = -\sum_{y \in Y} p(y) \ln(p(y))$$

[15]Per la proposizione 3.10 a pagina 10 la biezione $\varphi : Y \leftrightarrow X$ si estende ad un unico morfismo $\hat{\varphi} : Y^* \to A^*$.

In generale l'entropia di un sistema stocastico è una misura di quanto questo sistema sia differente da uno deterministico.

L'entropia verifica le seguenti proprietà:

(1) $H(S) \geq 0$;
   (a) $H(S) = 0 \iff \exists y_0 : p(y_0) = 1 \wedge p(y) = 0, \forall y \neq y_0$ (determinismo)

(2) $H(S) \leq \ln(n)$ con $n = Card(Y)$;
   (a) $H(S) = \ln(n) \iff \forall y \in Y, p(y) = \frac{1}{n}$ (tutte le lettere sono equiprobabili $\implies$ massimo indeterminismo)

Diseguaglianza di Gibbs.

PROPOSITION 5.8. *Siano p e q due distribuzione di Bernoulli su Y tali che q sia positiva. Allora si ha che*

$$\sum_{y \in Y} p(y) \ln \frac{1}{p(y)} \leq \sum_{y \in Y} p(y) \ln \frac{1}{q(y)}$$

L'uguglianza è verificata solo se $\forall y \in Y, p(y) = q(y)$.

DIMOSTRAZIONE. Supponiamo inizialmente che anche p sia positiva. Facendo la differenza tra primo e secondo membro della diseguaglianza, otteniamo

$$\sum_{y \in Y} p(y)(\ln \frac{1}{p(y)} - \ln \frac{1}{q(y)}) = \sum_{y \in Y} p(y)(\ln \frac{q(y)}{p(y)})$$

Dalle proprietà dei logaritmi sappiamo che $\ln(x) \leq x - 1$, quindi

$$\ln(\frac{q(y)}{p(y)}) \leq \frac{q(y)}{p(y)} - 1$$

Sostituendo quest'ultima all'interno della precedente, otteniamo

$$\sum_{y \in Y} p(y)(\ln \frac{q(y)}{p(y)}) \leq \sum_{y \in Y} p(y)(\frac{q(y)}{p(y)} - 1) = \sum_{y \in Y} (q(y) - p(y)) = 1 - 1 = 0$$

quindi

$$\sum_{y \in Y} p(y)(\ln \frac{q(y)}{p(y)}) \leq 0$$

L'uguaglianza si verifica quando

$$\sum_{y \in Y} p(y)(\ln \frac{q(y)}{p(y)}) = \sum_{y \in Y} p(y)(\frac{q(y)}{p(y)} - 1)$$

e questo accade se e solo se $\forall y \in Y, q(y) = p(y)$.

Ora vediamo cosa accade se non si fanno ipotesi restrittive su p. Innanzitutto dobbiamo effettuare le somme per quei valori di y tali che $p(y) \neq 0$ e per comodità indicheremo $\overline{\sum}$ la sommatoria che esclude automaticamente tali valori, ovvero $\overline{\sum_{y \in Y}} \equiv \sum_{y \in Y : p(y) \neq 0}$. A questo punto procediamo come in precedenza, effettuando al differenza tra le due somme

$$\overline{\sum_{y\in Y}} p(y)\ln\frac{1}{p(y)} - \overline{\sum_{y\in Y}} p(y)\ln\frac{1}{q(y)} = \overline{\sum_{y\in Y}} p(y)\ln\frac{q(y)}{p(y)} \leq \overline{\sum_{y\in Y}}(q(y)-p(y)) \leq 0$$

L'ultimo passaggio è giustificato da quanto segue.

Essendo $p \in DB$ e $q \in DBP$, per entrambe deve valere[16] che $\sum_{y\in Y} p(y) = 1$ e $\sum_{y\in Y} q(y) = 1$. Per quanto riguarda p vale ancora $\overline{\sum}_{y\in Y} p(y) = 1$ poiché gli elementi y esclusi sono quelli per cui $p(y) = 0$ e pertanto non inficiano il risultato della somma; il discorso è diverso per q dove $\overline{\sum}_{y\in Y} q(y) \leq 1$, dato che q è una distribuzione di Bernoulli positiva e l'esclusione dei già citati elementi inficia il risultato finale della somma.

Si conclude quindi che

$$\sum_{y\in Y} p(y)(\ln\frac{1}{p(y)}) \leq \sum_{y\in Y} p(y)(\ln\frac{1}{q(y)})$$

e, ancora una volta, l'uguaglianza si verifica se e soltanto se $\forall y \in Y, p(y) = q(y)$. □

Proprietà dell'entropia. Abbiamo già enunciato precedentemente le proprietà dell'entropia; ci accingeremo ora alla loro dimostrazione.

PROPOSITION 5.9. *$H(S) \geq 0$.*

DIMOSTRAZIONE. $H(S) = 0 \iff \sum_{y\in Y} p(y)(\ln\frac{1}{p(y)}) = 0 \iff p(y) = 0 \vee p(y) = 1$. Essendo $p \in DB$ allora $\sum_{y\in Y} p(y) = 1$ e quindi $\exists \overline{y} \in Y : p(\overline{y}) = 1$ e $\forall y \in Y : y \neq \overline{y}, p(y) = 0$; in tutti gli altri casi $\forall y \in Y, p(y) < 1 \implies H(S) > 0$. □

PROPOSITION 5.10. *$H(S) \leq \ln(n)$, con $n = Card(Y)$.*

DIMOSTRAZIONE. Scegliamo $q \in DBP$ tale che $\forall y \in Y, q(y) = \frac{1}{n}$. Per la diseguaglianza di Gibbs otteniamo

$$H(S) = \sum_{y\in Y} p(y)\ln\frac{1}{p(y)} \leq \sum_{y\ \in Y} \ln\frac{1}{1/n} = \sum_{y\in Y} p(y)\ln n = \ln n \sum_{y\in Y} p(y) = \ln n$$

quindi $H(S) \leq \ln n$. □

THEOREM 5.11. (**Teorema di Shannon**). *Sia $S = [Y,p]$ una sorgente e sia X un codice su A adattato a S. Detto φ il morfismo di codifica, allora*

$$C(X,\varphi) \geq \frac{H(S)}{\ln d}$$

con $d = Card(A)$

Potendo sempre scegliere φ tale che $C(X,\varphi) = C(X)$ allora, equivalentemente, vale anche

$$C(X) \geq \frac{H(S)}{\ln d}$$

[16]Vedere definizione 4.19 a pagina 22

Il valore minimo è raggiunto se e solo se il codice X è massimale e in tal caso $\forall y \in Y, p(y) = d^{-|\varphi(y)|}$.

DIMOSTRAZIONE. Sia $q \in DBP$ tale che

$$q(y) = \frac{d^{-|\varphi(y)|}}{\sum_{y \in Y} d^{-|\varphi(y)|}}$$

con $q(y) \in DB$.

Il denominatore, come si può notare, non dipende dal parametro, ma è costante e lo poniamo, pertanto, uguale ad α

$$\alpha = \sum_{y \in Y} d^{-|\varphi(y)|}$$

e quindi riscrivo $q(y)$ come segue

$$q(y) = \frac{d^{-|\varphi(y)|}}{\alpha}$$

Per la diseguaglianza di Gibbs,

$$H(S) = \sum_{y \in Y} p(y) \ln \frac{1}{p(y)} \leq \sum_{y \in Y} p(y) \ln \frac{1}{q(y)}$$

ma avendo definito $q(y)$, la diseguaglianza sopra diventa

$$\begin{aligned}
H(S) &= \sum_{y \in Y} p(y) \ln \frac{1}{p(y)} \leq \sum_{y \in Y} p(y) \ln \frac{\alpha}{d^{-|\varphi(y)|}} = \\
&= \sum_{y \in Y} p(y)(\ln \alpha - \ln d^{-|\varphi(y)|} = \\
&= \sum_{y \in Y} p(y)(\ln \alpha + \mid \varphi(y) \mid \ln d) =_{per\, \alpha\, e\, \ln d\, costanti} \\
&= \ln \alpha \sum_{y \in Y} p(y) + \ln d \sum_{y \in Y} \mid \varphi(y) \mid p(y)
\end{aligned}$$

Per definizione

$$C(X, \varphi) = \sum_{y \in Y} \mid \varphi(y) \mid p(y)$$

mentre come abbiamo visto

$$\sum_{y \in Y} p(y) = 1$$

quindi

$$\begin{aligned}
H(S) &\leq \ln \alpha + \ln d \cdot C(X, \varphi) \implies \\
&\implies \frac{H(S)}{\ln d} \leq \frac{\ln \alpha}{\ln d} + C(X, \varphi) \implies \\
&\implies \frac{H(S)}{\ln d} - \frac{\ln \alpha}{\ln d} \leq C(X, \varphi)
\end{aligned}$$

Poiché $\alpha = \sum_{y \in Y} d^{-|\varphi(y)|} = \sum_{y \in Y} d^{-|x|} \leq_{per\, dis.\, kraft-mcmillan} 1$, allora $\ln d \leq 0$; questo implica a sua volta che $-\frac{\ln \alpha}{\ln d} \geq 0$ e quindi $\frac{H(S)}{\ln d} \leq C(X, \varphi)$.

Per avere l'ugualianza deve accadere che $p \equiv q$ e $\ln \alpha = 0 \iff \alpha = 1 \iff \rho(X) = 1 \iff X$ è massimale. □

Il teorema di Shannon dimostra che la funzione costo ha un valore minimo in $\frac{H(S)}{\ln d}$ che viene raggiunto se e solo se il codice è massimale. In più ci fornisce la forma della distribuzione di probabilità.

Definition 5.12. Sia $X \subseteq A^+$, allora
X è un codice prefisso se $X \cap XA^+ = \oslash$
X è un codice suffisso se $X \cap A^+X = \oslash$
X è un codice biprefisso se è sia prefisso che suffisso
X è un codice uniforme se $\exists k > 0 : X \subseteq A^k$

Definition 5.13. X codice prefisso $\implies X^*$ unitario a sinistra.

Definition 5.14. (**Ordinamento prefissale**). Sia A finito, allora
$u \leq_p v$ (prefisso di) se $\exists h \in A^* : v = uh$ (oppure $v \in uA^*$ oppure $vA^* \subseteq uA^*$)
$u \leq_s v$ (suffisso di) se $\exists h \in A^* : v = hu \iff v \in A^*u$
$u \leq_f v$ (fattore di) se $\exists h, k \in A^* : v = huk \iff v \in A^*uA^*$
Tali ordinamenti sono relazioni d'ordine parziale; inoltre $\leq_p$ è transitiva.

Definition 5.15. (**Relazione di ricoprimento**). $u <_p v$ se $u \leq_p v$ e $\neg\exists w : u <_p w <_p v$.

Definition 5.16. (**Alberi di rappresentazione**). Sia A un alfabeto di di cardinalità d. Si può usare un albero d-ario che rappresenta la relazione di ricoprimento:
L'altezza del nodo rappresenta la lunghezza della parola.

Proposition 5.17. *Sia X un codice su A e sia T_X l'albero che lo rappresenta. X è prefisso se e soltanto se i nodi che lo rappresentano sono foglie.*

Definition 5.18. Un albero si dice completo se il grado dei suoi nodi interni è uguale a $d = Card(A)$, con A alfabeto.

Definition 5.19. Sia X un codice prefisso su A. X è un codice (prefisso) massimale se per ogni Y che sia codice massimale su A si ha che $X \subseteq Y \implies X = Y$.

Proposition 5.20. *Sia X un codice prefisso su A. Le seguenti affermazioni sono equivalenti:*
1) *X è un codice prefisso massimale*
2) $\forall w \in A^*\ wA^* \cap XA^* \neq \oslash$

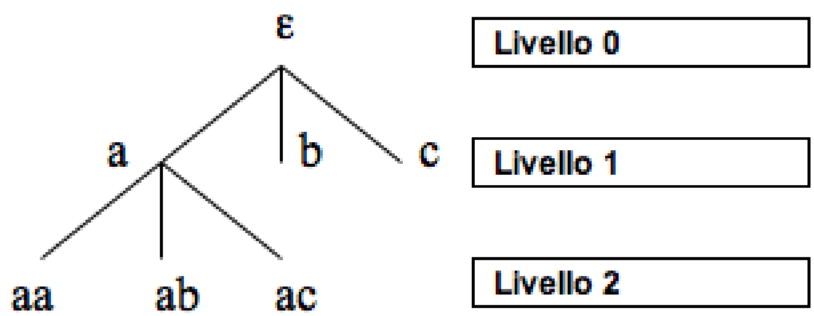

FIGURA 2. Albero di rappresentazione

3) $\forall f \in A^* \; fA^* \cap X^* \neq \oslash$
4) *L'albero che rappresenta X è completo*

DIMOSTRAZIONE. ($1 \implies 2$) Supponiamo per assurdo che esiste w tale che $wA^* \cap XA^* = \oslash$, allora l'insieme $Y = X \cup \{w\}$ è un codice prefisso e $X \subset Y$, che è contraddice l'ipotesi di X come codice massimale.

($2 \implies 3$) Ciò che ci accingiamo a fare è dimostrare che X è completo a destra. Poiché vale la 2), preso $f \in A^*$ avremo che $\exists u, v \in A^* \exists x \in X$ tali che $fu = xv$. Applicando il Lemma di Levi, si possono verificare due casi:

a) $x = ff_1 \implies f^*A^* \cap X^* \neq \oslash$

b) $f = xf_1$

Nel caso b bisogna applicare ricorsivamente il ragionamento. Proseguendo osserviamo ad ottenere una successione di parole $f, f_1, f_2, ..., f_i, ...$ dove $\mid f \mid > \mid f_1 \mid > ... > \mid f_i \mid$. Quindi $\exists i : \mid f_i \mid < \min\{\mid x \mid : x \in X\}$. Qui la derivazione destra non può continuare e avremo:

($3 \implies 1$) Supponiamo per assurdo che X non sia massimale. Questo implica che $\exists f : X \cup \{f\} = Y$ è codice.

Per la 3) allora esiste $p \in A^*$ tale che $fp \in X^*$ ovvero $fp = x_1...x_n$ con $x_i \in X$. Per il lemma di Levi possiamo avere due casi: o f prefisso di x_1 o x_1 è prefisso di f. In entrambi i casi si otterebbe che $X \cup \{f\}$ non può essere prefisso, che è assurdo.

($1 \implies 4$) Per ipotesi X è prefisso massimale. Supponiamo per assurdo che l'alfabeto che lo rappresenta non sia completo. Allora costruiamo un nuovo codice aggiungendo il ramo che manca. Il nuovo codice includerà il precedente e sarà prefisso; questo implicherebbe che X non è prefisso massimale, raggiungendo un assurdo.

($4 \implies 2$) Se l'albero è completo, l'unica possibilità per wA^* è che si trovi su un ramo o nel sottoalbero generato da un nodo. □

DEFINITION 5.21. Un codice prefisso finito è finitamente completabile.

PROPOSITION 5.22. *Sia X un codice prefisso non denso. Le seguenti condizioni sono equivalenti:*

1) *X è massimale come prefisso*

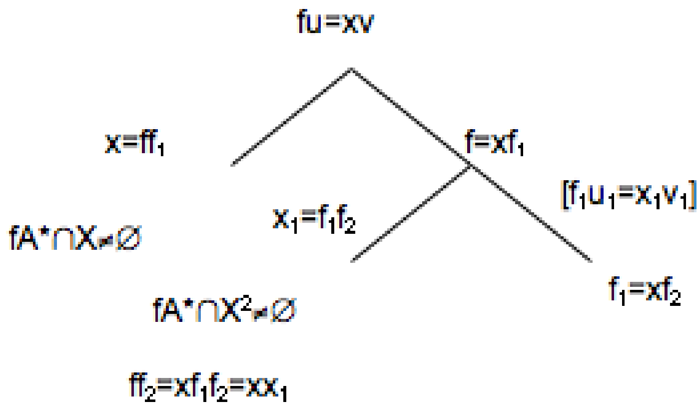

FIGURA 3

2) X *è completo a destra* ($\forall f \in A^*$ $fA^* \cap X^* \neq \oslash$)
3) X *è completo* ($\forall f \in A^*$ $A^*fA^* \cap X^* \neq \oslash$)
4) $\forall \mu \in DBP$ $\mu(X) = 1$
5) $\exists \mu \in DBP : \mu(X) = 1$
6) X *è massimale come codice*

DIMOSTRAZIONE. ($1 \implies 2$) Già dimostrato nella proposizione 5.20 a pagina 36.

($2 \implies 3$) Banale. ($A^* \supseteq \{\varepsilon\}$, $\varepsilon f w = x_1 ... x_n$ e w esiste per la 2))

($3 \implies 4$) Per il teorema di Kraft-McMillan (4.30 a pagina 24), essendo X codice si ha che $\mu(X) \leq 1$. Dalla proposizione 4.38 a pagina 26, essendo X codice non denso e completo si ha che $\mu(X) \geq 1$. Ergo, $\forall \mu \in DBP$ $\mu(X) = 1$.

($4 \implies 5$) Banale.

($5 \implies 6$) Per la proposizione 4.32 a pagina 26.

($6 \implies 1$) Se X è massimale come codice allora X è massimale come prefisso. □

THEOREM 5.23. (**Teorema di Kraft**). *Sia* $f : \mathbb{N} \to \mathbb{N}$ *una qualsiasi funzione tale che* $f(0) = 0$ *e* $\sum_{n\geq 0} f(n)d^{-n} \leq 1$ *con* d *un intero maggiore di zero. Allora esiste sempre un codice prefisso* X *su un alfabeto* A *di cardinalità* d *con funzione di struttura* f*, ovvero* $f_X \equiv f$.

DIMOSTRAZIONE. Sappiamo che $\sum_{n\geq 0} f(n)d^{-n} \leq 1$, quindi fissato $\Delta > 0$ possiamo scrivere

$$\sum_{n=0}^{\Delta} f(n)d^{-n} \leq 1 \implies f(\Delta)d^{-\Delta} + \sum_{n=0}^{\Delta-1} f(n)d^{-n} \leq 1 \implies$$

$$\implies f(\Delta)d^{-\Delta} \leq 1 - \sum_{n=0}^{\Delta-1} f(n)d^{-n} \implies f(\Delta) \leq d^{\Delta} - \sum_{n=0}^{\Delta-1} f(n)d^{\Delta-n}$$

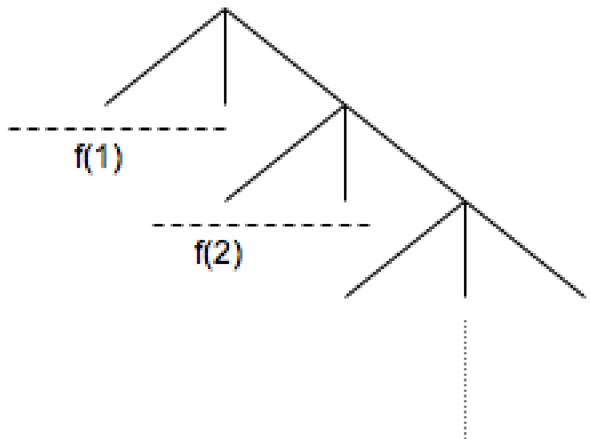

FIGURA 4

Poniamo

$$v(\Delta) = d^{\Delta} - \sum_{n=0}^{\Delta-1} f(n) d^{\Delta-n}$$

e scriviamo più compattamente

$$f(\Delta) \leq v(\Delta)$$

Dimostriamo ora che $v(\Delta+1) = d[v(\Delta) - f(\Delta)]$.

$$\begin{aligned} v(\Delta+1) &= d^{\Delta+1} - \sum_{n=0}^{\Delta} f(n) d^{\Delta+1-n} = d(d^{\Delta} - \sum_{n=0}^{\Delta} f(n) d^{\Delta-n}) = \\ &= d(d^{\Delta} - \sum_{n=0}^{\Delta-1} f(n) d^{\Delta-n} - f(\Delta) d^{\Delta-\Delta}) = d(v(\Delta) - f(\Delta)) \end{aligned}$$

Adesso consideriamo l'albero d-ario per costruire il codice prefisso X tale che $f_X = f$. Sappiamo che $f(1) \leq v(1) = d$. Possiamo scegliere allora di potare $f(1)$ nodi. Dai nodi rimanenti $(v(1) - f(1))$ si dipartono $d(v(1) - f(1)) = v(2)$ figli. Di questi ne possiamo potare $f(2) \leq v(2)$. I nodi che restano a un livello sono prefissi per i nodi del livello successivo. Iterando il processo si costruisce un albero che rappresenta un codice prefisso avente la seguente struttura:

$f_X(1) = f(1)$
$f_X(2) = f(2)$
...

□

Il teorema di Kraft ci dice che data una qualsiasi funzione tale che soddisfi la diseguaglianza di Kraft-McMillan (4 a pagina 23), è possibile trovare un codice prefisso che abbia proprio quella funzione come funzione di struttura.

COROLLARY 5.24. *Sia X un codice su un alfabeto A di cardinalità $d > 0$. Allora esiste un codice prefisso Y su A tale che ha la stessa distribuzione delle lunghezze ($f_X \equiv f_Y$).*

DIMOSTRAZIONE. Se X è un codice su A e $d = Card(A)$, allora per la diseguaglianza di Kraft-McMillan abbiamo

$$\sum_{n \geq 0} f_X(n) d^{-n} \leq 1$$

e $f_X(0) = 0$ poiché $\varepsilon \notin X$. Pertanto, in seguito al teorema di Kraft (5.23 a pagina 38), esiste un codice prefisso Y tale che $f_Y \equiv f_X$ □

PROPOSITION 5.25. *Sia $S = [Y, p]$ una sorgente di informazione. Per ogni codice X su A con $d = card(A)$ esiste sempre un codice prefisso Z tale che $C(X) = C(Z)$.*

DIMOSTRAZIONE. Per il corollario 5.24 esiste Z tale che $f_X = f_Z$. Si può allora costruire un morfismo $\varphi : X \to Z$ in modo che

$$| x |=| \varphi(x) |$$

con $\varphi(x) \in Z$
Per definizione di costo abbiamo

$$C(X) = \sum_{x \in X} p(x) \mid x \mid= \sum_{x \in X} p(x) \mid \varphi(x) \mid= \sum_{z \in \varphi(X)} p(z) \mid z \mid= C(Z)$$

□

COROLLARY 5.26. *Esiste sempre un codice prefisso ottimale per la sorgente $S = [Y, p]$.*

PROPOSITION 5.27. *Sia X un codice su un alfabeto a due lettere e sia esso ottimale per la sorgente $S = [Y, p]$. Allora X è massimale.*

DIMOSTRAZIONE. Per la proposizione 5.25, esiste un codice Y_2 prefisso, il cui costo $C(Y_2) = C(X)$ e quindi Y_2 è ottimale. L'albero rappresentante Y_2 è completo. Se per assurdo l'albero non fosse completo, essendo un albero binario, esisterebbe un nodo con un solo figlio. Tale nodo può essere eliminato e si otterrebbe un codice con un costo minore, il che porterebbe a un assurdo, poiché Y_2 è ottimale.

Essendo l'albero completo, inoltre, per la proposizione 5.20 a pagina 36, Y_2 è prefisso massimale. In aggiunta, Y_2 ha la stessa distribuzione delle lunghezze di X, ed essendo Y_2 massimale, si ha:

$$\sum_{n \geq 0} f_{Y_2}(n) d^{-n} = 1 \implies \sum_{n \geq 0} f_X(n) d^{-n} = 1 \implies X \text{ } massimale$$

□

PROPOSITION 5.28. *Sia $S = [Y, p]$ una sorgente di informazione. Allora esiste sempre un codice prefisso Z su A tale che*

$$\frac{H(S)}{\ln d} \leq C(Z) < \frac{H(S)}{\ln d} + 1$$

con $d = Card(A)$.

DIMOSTRAZIONE. Dal teorema di Shannon, ricordiamo che se si è raggiunto $\frac{H(S)}{\ln d}$ allora deve essere $p(y) = d^{-|\varphi(y)|}$ quindi $\log_d p(y) = - \mid \varphi(y) \mid$ cioè $\mid \varphi(y) \mid = \log_d \frac{1}{p(y)}$. Consideriamo $l(y) = \lceil \log_d \frac{1}{p(y)} \rceil$ che è il più piccolo intero maggiore o uguale di $\log_d \frac{1}{p(y)}$; possiamo quindi scrivere

$$\log_d \frac{1}{p(y)} \leq l(y) < \log_d \frac{1}{p(y)} + 1$$

Allora si avrà anche che

$$\begin{aligned} -\log_d \frac{1}{p(y)} \quad &\geq \quad -l(y) > -\log_d \frac{1}{p(y)} - 1 \implies d^{-\log_d \frac{1}{p(y)}} \geq d^{-l(y)} \implies \\ &\implies \quad p(y) \geq d^{-l(y)} \ \forall y \in Y \implies \sum_{y \in Y} d^{-l(y)} \leq \sum_{y \in Y} p(y) = 1 \end{aligned}$$

Rifacendoci al teorema di Kraft, possiamo scrivere $\sum_{y \in Y} d^{-l(y)}$ come la somma di una lunghezza per d^{-n}, quindi esiste un codice prefisso Z che ha parole di lunghezza $l(y_1), ..., l(y_n)$.

Torniamo ora alla disequazione

$$\log_d \frac{1}{p(y)} \leq l(y) < \log_d \frac{1}{p(y)} + 1$$

e moltiplichiamo per $p(y)$ ottenendo

$$p(y) \log_d \frac{1}{p(y)} \leq p(y) l(y) < p(y) \log_d \frac{1}{p(y)} + p(y)$$

Se sommiamo rispetto a y otteniamo

$$\sum_{y \in Y} p(y) \log_d \frac{1}{p(y)} \leq \sum_{y \in Y} p(y) l(y) < \sum_{y \in Y} p(y) \log_d \frac{1}{p(y)} + \sum_{y \in Y} p(y)$$

ovvero

$$\frac{H(S)}{\ln d} \leq C(Z, \hat{\varphi}) < \frac{H(S)}{\ln d} + 1$$

con $\hat{\varphi} : y \to l(y)$.

Se consideriamo il minimo $C(Z)$ otteniamo $C(Z) < \frac{H(S)}{\ln d} + 1$. □

DEFINITION 5.29. Data $S = [Y, p]$ una sorgente, associamo ad essa $S_n = [Y^n, p_n]$, detta sorgente estesa di ordine n, con $Y^n = \underbrace{YY...Y}_{n\,volte}$ e $p_n(l) = p(y_1)...p(y_n)$ tale che $l \in Y^n, l = y_1 y_2 ... y_n$.

LEMMA 5.30. $\forall n > 0$ *si ha che* $H(S_n) = nH(S)$

DIMOSTRAZIONE. La dimostrazione segue per induzione su n.
Passo base. $n = 1 \implies H(S_1) = H(S)$ in quanto $S_1 = S$
Passo di induzione.

$$H(S_n) = \sum_{z \in Y^n} p(z) \ln \frac{1}{p(z)}$$

Poiché $z \in Y^n \implies z = ty$ con $t \in Y^{n-1}$ e $y \in Y$. Quindi

$$\begin{aligned} H(S_n) &= \sum_{t \in Y^{n-1}, y \in Y} p(t)p(y) \ln(p(t)p(y))^{-1} = -\sum_{t \in Y^{n-1}, y \in Y} p(t)p(y)[\ln p(t) + \ln p(y)] = \\ &= -\sum_{t \in Y^{n-1}, y \in Y} p(t)p(y) \ln p(t) - \sum_{t \in Y^{n-1}, y \in Y} p(t)p(y) \ln p(y) = \\ &= \sum_{y \in Y} p(y) \sum_{t \in Y^{n-1}} p(t) \ln p(t) - \sum_{y \in Y} p(y) \ln p(y) \sum_{t \in Y^{n-1}} p(t) \end{aligned}$$

Dato che $\sum_{y \in Y} p(y) = 1 = p(Y)$ e $\sum_{t \in Y^{n-1}} p(t) = 1 = p(Y^{n-1})$ allora si ha

$$H(S_{n-1}) + H(S) =_{per\, ipotesi\, induttiva} (n-1)H(S) + H(S) = nH(S)$$

□

Costruzione di un codice prefisso ottimale: Metodo di Huffman. Il metodo di Huffman riguarda codici su alfabeti a due simboli ($A = \{0, 1\}$).

Il metodo di Huffman consente di calcolare, a partire da una sorgente, un codice prefisso ottimale. Ricordiamo che il teorema di Shannon ci fornisce un limite inferiore per il costo di un codice. Intuitivamente il metodo di Huffman associa alle lettere meno probabili un codice più lungo e un codice più breve alle lettere più probabili. Il metodo di Huffman consiste nella creazione di una successione di sorgenti nella quale si passa dalla sorgente S_i alla sorgente S_{i+1} collassando in un unico simbolo i due simboli (elementi dell'alfabeto) a cui è associata la probabilità ($p(y)$) minore. La probabilità del nuovo simbolo introdotto sarà data dalla somma delle probabilità dei due simboli collassati. Il procedimento si itera finché non si arriva a una sorgente avente due soli simboli.

EXAMPLE.

SIMB	P	SIMB	P	SIMB	P	SIMB	P	SIMB	P
y_1	0.3	y_1	0.3	y_1	0.3	$y_{2,3}$	0.45	$y_{1,4,5,6}$	0.55
y_2	0.25	y_2	0.25	$y_{4,5,6}$	0.25	y_1	0.3	$y_{2,3}$	0.45
y_3	0.2	y_3	0.2	y_2	0.25	$y_{4,5,6}$	0.25		
y_4	0.1	$y_{5,6}$	0.15	y_3	0.2				
y_5	0.1	y_4	0.1						
y_6	0.05								

Per trovare il codice associato si procede in maniera inversa associando i due simboli dell'alfabeto ai simboli dell'ultima sorgente trovata e man mano che si procede a ritroso ad ogni simbolo aggregato che si divide.

SIMB	COD	SIMB	COD	SIMB	COD	SIMB	COD	SIMB	COD
$y_{1,4,5,6}$	0	$y_{2,3}$	1	y_1	00	y_1	00	y_1	00
$y_{2,3}$	1	y_1	00	$y_{4,5,6}$	01	y_2	10	y_2	10
		$y_{4,5,6}$	01	y_2	10	y_3	11	y_3	11
				y_3	11	$y_{5,6}$	010	y_4	011
						y_4	011	y_5	0100
								y_6	0101

Dimostrazione della correttezza di Huffman. Sia $X = \{x_1, x_2, ..., x_{m-1}, x_m\}$ un codice prefisso ottimale; si ha

1) $\forall x, y \in X\ p(x) > p(y) \implies |\ x_1 \ |\leq|\ x_2\ |$

2) I due simboli con minore probabilità, x_{m-1} e x_m, hanno la stessa lunghezza $L = \max\{|\ x\ |: x \in X\}$ e possiedono inoltre lo stesso prefisso di lunghezza $L-1$.

Supponiamo di dover passare dalla sorgente S alla sorgente $S^{'}$

$S = \begin{pmatrix} y_1 & \dots & y_{m-1} & y_m \\ p_1 & \dots & p_{m-1} & p_m \\ x_1 & \dots & x_{m-1} & x_m \end{pmatrix}$, con $X = \{x_1, x_2, ..., x_{m-1}, x_m\}$, $Y = \{y_1, y_2, ..., y_{m-1}, y_m\}$ e $P = \{p_1, p_2, ..., p_{m-1}, p_m\}$

$S^{'} = \begin{pmatrix} y_1 & \dots & y_{m,m-1} & y_{m-2} \\ p_1 & \dots & p_m + p_{m-1} & p_{m-2} \\ x_1 & \dots & x_{m,m-1} & x_{m-2} \end{pmatrix}$, con $X^{'} = \{x_1, x_2, ..., x_{m,m-1}, x_{m-1}\}$, $Y^{'} = \{y_1, y_2, ..., y_{m,m-1}, y_{m-2}\}$ e $P^{'} = \{p_1, p_2, ..., p_{m-1} + p_m, p_{m-2}\}$

Il passaggio da $S^{'}$ a S si ottiene concatenando a x_m, x_{m-1} una volta 0 e una volta 1.

Vogliamo dismostrare che se $X^{'}$ è ottimale per $S^{'}$ allora X è ottimale per S.

Si considerino i costi di X e di $X^{'}$:

$$C(X) = \sum_{i}^{m} p_i |\ x_i\ |$$

$$C(X^{'}) = \sum_{i}^{m-2} p_i \mid x_i \mid + (p_m + p_{m-1}) \mid x_{m,m-1} \mid$$

A questo punto, supponiamo per assurdo che X non sia ottimale per S; allora esiste $\hat{X} = \{\hat{x}_1, \hat{x}_2, ..., \hat{x}_{m-1}, \hat{x}_m\}$ ottimale per S tale che $C(\hat{X}) < C(X)$. A partire da $\hat{X}$ possiamo costruire il codice $\hat{X}^{'}$ con la procedura di Huffman al contrario, togliendo una lettera da $\hat{x}_m$ e $\hat{x}_{m-1}$ e ottenendo così il prefisso comune di lunghezza $L-1$, $\hat{x}_{m,m-1}$; pertanto $\hat{X}^{'} = \{\hat{x}_1, \hat{x_2}, ..., \hat{x}_{m,m-1}, \hat{x}_{m-2}\}$.

La dimostrazione prosegue ora per arrivare a dimostrare che $C(\hat{X}^{'}) < C(X^{'})$, raggiungendo così un assurdo, dato che $X^{'}$ è supposto ottimale per $S^{'}$

$$C(\hat{X}^{'}) = \sum_{i=1}^{m-2} p_i \mid \hat{x}_i \mid + (p_m + p_{m-1}) \mid \hat{x}_{m,m-1} \mid$$

Ricordando che $\mid x^{'}_{m,m-1} \mid = \mid x^{'}_m \mid -1 = \mid x^{'}_{m-1} \mid -1$ otteniamo

$$\begin{aligned}
C(\hat{X}') &= \sum_{i=1}^{m-2} p_i \mid x_i' \mid + p_{m-1} \mid x_m' \mid + p_{m-1} \mid x_{m-1}' \mid - p_m - p_{m-1} = \\
&= \sum_{i=1}^{m} p_i \mid x_i' \mid - p_m - p_{m-1} = C(\hat{X}) - p_m - p_{m-1} < C(X) - p_m - p_{m-1} = \\
&= \sum_{i=1}^{m} p_i \mid x_i \mid - p_m - p_{m-1} = \sum_{i=1}^{m-2} p_i \mid x_i \mid + p_m \mid x_m \mid + p_{m-1} \mid x_{m-1} \mid - p_m - p_{m-1} = \\
&= \sum_{i=1}^{m-2} p_i \mid x_i \mid + p_m(\mid x_{m,m-1} \mid + 1) + p_{m-1}(\mid x_{m,m-1} \mid + 1) - p_m - p_{m-1} = \\
&= \sum_{i=1}^{m-2} p_i \mid x_i \mid + p_m \mid x_{m,m-1} \mid + p_m + p_{m-1} \mid x_{m,m-1} \mid + p_{m-1} - p_m - p_{m-1} = \\
&= \sum_{i=1}^{m-2} p_i \mid x_i \mid + (p_m + p_{m-1}) \mid x_{m,m-1} \mid = C(X')
\end{aligned}$$

Si è giunti dunque a dire che $C(\hat{X}') < C(X')$ raggiungendo il già citato assurdo e quindi X è ottimale per S.

Ridardo di decifrazione. In generale affinché un messaggio in codice possa essere decodificato è necessario che l'intero testo in codice sia trasmesso. In alcuni casi, per codici particolari, il ritardo di decifrazione può essere finito.

Definition 5.31. Sia X un insieme di parole su A e sia d un intero maggiore di zero. Allora X ha ritardo d se si verifica la proprietà $u(d)$

$$\forall x, x' \in X, xX^dA^* \cap x'X^* \neq \oslash \implies x = x'$$

Tale definizione può essere interpretata nel modo seguente: supponiamo di avere una sequenza di lettere da sinistra verso destra. Se riconosiamo una parola x del codice X, ci basta riconoscere al più d parole appartenenti al codice X dopodiché si è sicuri che la fattorizzazione della sequenza è univocamente determinata.

Proposition 5.32. *Se X verifica $u(d)$ allora X è codice.*

Dimostrazione. Supponiamo che X non sia un codice, allora $\exists x_1, x_1', y_1, ..., y_r, y_1', ..., y_s' \in X$ tali che $x_1y_1...y_r = x_1'y_1'...y_s'$ con $x_1 \neq x_1'$. Moltiplichiamo ambo i membri per $\underbrace{yy...y}_{d\,volte}$ (con $y \in X$) e otteniamo $x_1 \underbrace{y_1...y_ryy...y}_{d\,elementi\,di\,X} = x_1'y_1'...y_s'yy...y$. Si ha allora che

$$x_1X^d\xi \in x_1'X^*$$

poiché vale $u(d) \implies x_1 = x_1'$ che ci porta a contraddizione. $\square$

Proposition 5.33. *Se X verifica $u(d)$ allora X verifica $u(d+1)$.*

Proposition 5.34. *Un codice X è prefisso se e solo se X ha ritardo di decifrazione pari a zero.*

THEOREM 5.35. (**Teorema di Even**). *Sia X un codice finito, X ha ritardo di decifrazione finito se e solo se esiste $n > 0$ tale che $R_n = \oslash$*

Questo teorema ci conferisce la facoltà di sapere quando un codice finito ha ritardo di decifrazione finito o infinito.

THEOREM 5.36. (**Terzo Teorema di Schützenberger**). *Sia X un codice massimale finito, allora X è prefisso oppure ha ritardo di decifrazione infinito.*

COROLLARY 5.37. *Sia $S = [Y, p]$ una sorgente e sia X un codice su un alfabeto a due lettere ottimale per S. Allora X è prefisso oppure ha ritardo di decifrazione infinito.*

DEFINITION 5.38. Sia X un codice. La coppia $(u, v) \in X^+ x X^+$ è detta **coppia sincronizzante** per X^+e $\forall s, t \in A^*\ suvt \in X^* \implies su, vt \in X^*$.

DEFINITION 5.39. Un codice X si dice **sincronizzante** se esiste almeno una coppia sincronizzante per X^+.

PROPOSITION 5.40. *Sia X un codice su A massimale e sincronizzante. Allora $\exists c \in X^+ : cA^*c \subseteq X^*$.*

DIMOSTRAZIONE. X codice massimale $\implies \forall f \in A^*\ A^*fA^* \cap X^* \neq \oslash$, cioè ogni parola f si può completare. Esiste inoltre $c \in X^+$ tale che $cfc \in X^*$.

Sia $f \in A^*$ tale che $f = pqfpq$

X massimale $\implies X$ è completo $\implies \exists \lambda, \mu \in A^* : \lambda pqfpqu \in X^* \implies \lambda p, yfp, q\mu \in X^* \implies$

$\implies qfp \in X^* \implies \underset{c}{pq}f\underset{c}{pq} \in X^* \implies cfc \in X^* \implies cA^*c \subseteq X^*$. □

PROPOSITION 5.41. *Se X è un codice (non denso) tale che $\exists c \in X^+ : cA^*c \subseteq X^*$ allora X è massimale e sincronizzante.*

DIMOSTRAZIONE. Dobbiamo dimostrare che la coppia $(c, c) \in X^+ x X^+$ è sincronizzante per X^+. Supponiamo che $\lambda cc\mu \in X^*$ allora anche $\lambda cc\mu c \in X*$, $c\mu cc \in X^*$ e $c\lambda c \in X^*$, quindi ho che $\lambda c \in (X^*)^{-1}X \cap X^*(X^*)^{-1} \implies \lambda c \in X^*$.

Poiché $\lambda c \in X^*$ e $\begin{cases} \lambda cc\mu \in X^* \\ c\mu c \in X^* \end{cases}$ allora $c\mu \in X^*$ per il (terzo) teorema di Schützenberger; quindi la coppia $(\lambda, u) \in X^+ x X^+$ è sincronizzante per X^+. □

Conseguenza. Se X è un codice biprefisso, $X \neq A$ massimale $\implies X$ non è sincronizzante.

DIMOSTRAZIONE. Se X è sincronizzante e massimale, allora $\exists c \in X^+ : cA^*c \subseteq X^*$ cioè $\forall f \in A^*\ cfc \in X^*$.

Essendo X prefisso allora X^+ è unitario a sinistra, quindi $cfc \in X^* \implies fc \in X^*$. Ma X è anche suffisso, quindi unitario a destra e quindi $f \in X^* \implies A^* \subseteq X^* \subseteq A^* \implies X^* = A^* \implies X = A$ il che è assurdo. □

Bibliografia

[1] Alberto Facchini, *Algebra e Matematica discreta*, Zanichelli-Decibel 2000
[2] Rudolf Lidl, Günter Pilz, *Applied Abstract Algebra - Second Edition*, Springer-Verlag 1998
[3] *Entropia (Teoria dell'Informazione)*,
(http://it.wikipedia.org/wiki/Entropia_(teoria_dell'informazione))