UNIVERSITA' DEGLI STUDI DI PISA

ISTITUTO DI SCIENZE DELL'INFORMAZIONE

Introduzione alle sorgenti linguistiche binarie

Sergio Frasca

NOTE SCIENTIFICHE

S 74 16

NOVEMBRE 1974

Università degli Studi di Pisa
ISTITUTO DI SCIENZE DELL'INFORMAZIONE

INTRODUZIONE ALLE SORGENTI LINGUISTICHE BINARIE*

Sergio Frasca**

Nota Scientifica S-74-16
(Novembre 1974)

---

* Lavoro eseguito nell'ambito del GNAFA-CNR.

** Borsista presso l'Istituto di Scienze dell'In

RIASSUNTO

Si presenta un nuovo modello matematico di sorgente binaria. Tale modello è più adeguato ai casi pratici del modello markoviano perchè

- tiene conto delle dipendenze tra i bit a qualsiasi distanza (e non solo fino a una distanza finita)
- tiene conto che tali dipendenze decrescono al crescere della distanza tra i bit.

E' presentato un mezzo di studio (le trasformate di Walsh) particolarmente adeguato al modello e sono ricavati alcuni interessanti risultati preliminari.

1 - SEQUENZE BINARIE MARKOVIANE.

Si intende per "sequenza markoviana binaria di ordine ω" una sequenza $A \equiv \dots a_{-k}\, a_{-k+1} \dots a_{-\omega}\, a_{-\omega+1} \dots a_{-2}\, a_{-1}$ (che supporremo di lunghezza infinita) i cui elementi appartengano all'insieme $\{0,1\}$ e in cui la probabilità che l'elemento (o bit) a_i sia 1 (o 0) dipende dal valore degli ω precedenti $a_{i-\omega}, a_{i-\omega+1}, \dots a_{i-1}$. Più precisamente, chiamando S_i (o "stato i-esimo) l'i-esima ω-upla di 1 e 0[1] possiamo vedere la sequenza A come successione non più di bit, ma di stati, cioè di elementi dell'insieme $\{S_0, S_1, \dots, S_{2^\omega-1}\}$. Tale successione è, ovviamente, una catena di Markov semplice e la cui matrice stocastica è

$$T = \begin{matrix}
q_0 & p_0 & 0 & 0 & 0 & 0 & 0 \;.\;.\;. & 0 & 0 \\
0 & 0 & q_1 & p_1 & 0 & 0 & 0 \;.\;.\;. & 0 & 0 \\
0 & 0 & 0 & 0 & q_2 & p_2 & 0 \;.\;.\;. & 0 & 0 \\
. & . & . & . & . & . & .\;.\;.\;.\;. & . & . \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \;.\;.\;. & q_{2^{\omega-1}-1} & p_{2^{\omega-1}-1} \\
q_{2^{\omega-1}} & p_{2^{\omega-1}} & 0 & 0 & 0 & 0 & 0 \;.\;.\;. & 0 & 0 \\
0 & 0 & q_{2^{\omega-1}+1} & p_{2^{\omega-1}} & 0 & 0 & 0 \;.\;.\;. & 0 & 0 \\
. & . & . & . & . & . & .\;.\;.\;.\;. & . & . \\
0 \quad 0 & 0 & 0 & 0 & 0 & 0 & 0 \;.\;.\;. & q_{2^\omega-1} & p_{2^\omega-1}
\end{matrix}$$

Come si vede gli elementi della matrice $T = \{t_{ij}\}$ sono tutti nulli meno quelli della forma $q_i = t_{i,2i}$ e $p_i = t_{i,2i+1}$, per $i < 2^{\omega-1}$ e della forma $q_i = t_{i-2^{\omega-1},2(i-2^{\omega-1})}$ e $p_i = t_{i-2^{\omega-1},2(i-2^{\omega-1})+1}$, per $i \geq 2^{\omega-1}$, inoltre $q_1 = 1 - p_i$.

[Nota (1): cioè

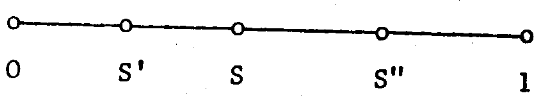
	$a_{-\omega}$	$a_{-\omega+1}$		a_{-3}	a_{-2}	a_{-1}
S_0 =	0	0	. . .	0	0	0
S_1 =	0	0	. . .	0	0	1
S_2 =	0	0	. . .	0	1	0
.	. .		. . . . .		.	.
$S_{2^\omega-1}$ =	1	1	. . .	1	1	1]

Si noti che i p_i sono relativi a transizioni verso stati di ordine dispari nella sequenza degli stati e cioè a transizioni verso l'1 nella sequenza binaria, e viceversa i q_i sono relativi a transizioni verso stati di ordine pari e cioè, nella sequenza binaria, verso lo 0. Nel caso in cui la catena sia ergodica si può definire la probabilità di presenza degli stati. Chiameremo P_i la probabilità di presenza dello stato S_i; ricordiamo che essa è l'i-esima componente dell'autovettore della matrice T relativo all'autovalore 1 (che, in caso di ergodicità, esiste sempre ed è unico).

2 - SEGMENTO DELLE SEQUENZE.

Introduciamo ora un nuovo modo di descrivere una catena di Markov binaria, meno "ingormbrante" della matrice di transizione; esso si mostrerà molto utile nel seguito.

Consideriamo un segmento di lunghezza unitaria ed esprimiamo in numerazione binaria l'ascissa dei suoi punti. In tal modo è possibile far corrispondere ad ogni punto del segmento una sequenza infinita di 1 e 0. Ciò ci suggerisce di adoperare tale segmento per rappresentare tutte le possibili sequenze binarie, con la convenzione che le cifre più significative siano le più recenti:



Si vede subito che, se si parte della sequenza corrispondente al punto S, essa può, al passo successivo, trasformarsi in quelle corrispondenti ai due punti

$S' = \frac{S}{2}$ se si ha una transizione verso lo 0

$S'' = \frac{S+1}{2}$ se si ha una transizione verso l'1

Tali due punti sono i centri dei due segmenti staccati da S sul segmento unitario e si ha che

$$S'' = S' + 0.5.$$

Definiamo ora sul nostro segmento la funzione p(S) che indica la probabilità che, essendo finora uscita la sequenza S, si abbia, al prossimo passo, 1. Nel caso di catene di Markov di ordine ω tale probabilità è la stessa per tutte le sequenze che hanno uguali gli ultimi ω bit e sono quindi nello stesso stato. Tali sequenze sono rappresentate sul segmento unitario dai punti di un suo opportuno intervallo di ampiezza $2^{-\omega}$: per lo stato i-esimo, tale intervallo è quello definito per tutti gli x per cui $i \cdot 2^{-\omega} \leq x \leq (i+1) . 2^{-\omega}$. La p(S) avrà quindi in genere un numero finito di discontinuità (2^{ω} -1 al più) in punti multipli di $2^{-\omega}$ e tra due successivi punti di discontinuità sarà costante. Definiremo in seguito la forma più generale per la p(S).

3 - TRASFORMATE DI WALSH.

Introdurremo ora uno strumento matematico molto utile in questo approccio allo studio delle sequenze binarie: le funzioni e le trasformate di Walsh.

Con il nome di "funzioni di Walsh" si indica in genere un insieme di funzioni ortonormali completo che hanno la particolarità di poter assumere solo due valori: 1 e -1. Esistono molti modi di definirle e in genere anche il modo di ordinarle non è unico, come il dominio in cui di solito vengono definite. Taluni le hanno definite in $(-\frac{1}{2},\frac{1}{2})$, le hanno divise in sal (funzioni dispari) e cal (funzioni pari) in analogia alle funzioni seno e coseno e hanno per esse definito una "frequenza generalizzata" o "sequenza" che indica il numero di volte che cambiano di segno nel dominio. Noi, per i nostri fini, usiamo la seguente definizione.

Le funzioni di Walsh $\{W_i\}$ sono un insieme di funzioni definito nell'intervallo (0,1) dell'asse reale, che godono delle seguenti proprietà:

1) prendono i soli due valori 1 e -1

2) sono ortonormali tra di loro, cioè

$$\int_0^1 W_i(S) * W_j(S) \, d\, S = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i = j \end{cases}$$

3) costituiscono un insieme completo, cioè, comunque data una $f(S)$ quadrato sommabile in $(0, 1)$ si può scrivere, essendo k_i dei coefficienti opportuni indipendenti da S,

$$f(S) = \Sigma_i k_i * W_i(S)$$

4) la funzione $W_i(S)$ ha "sequenza" i, cioè cambia di segno i volte in (0,I).

Si può dimostrare che tale insieme è unico, a parte inessenziali traslazioni e cambiamenti di segno.

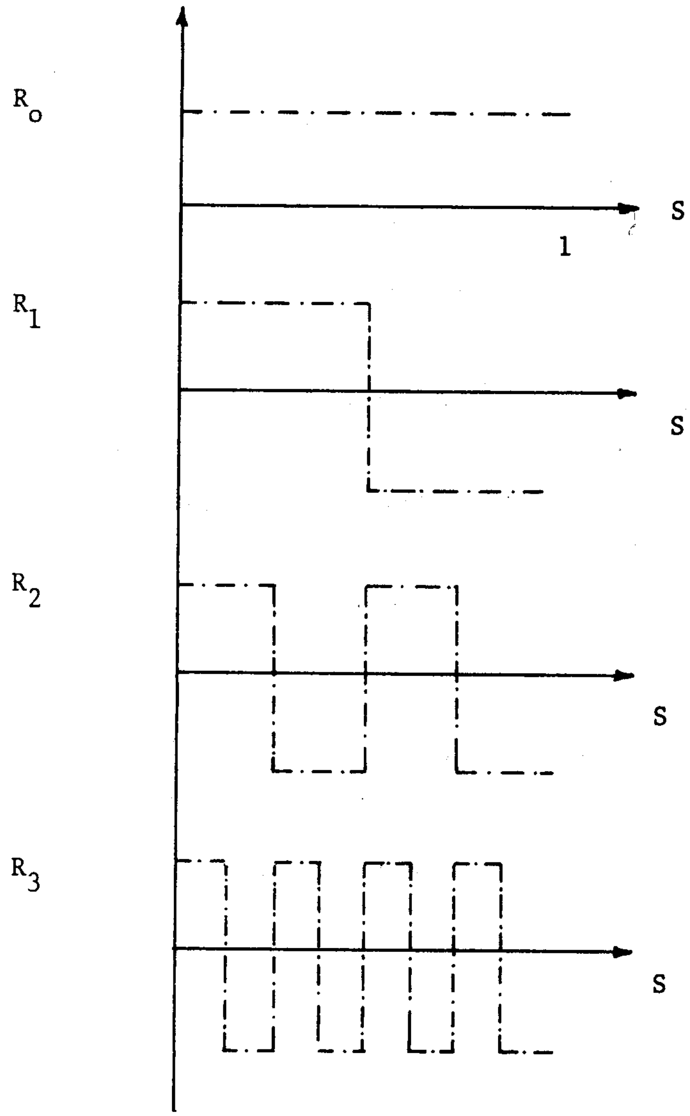
Per generarle utilizziamo il procedimento di LACKEY e MELTZER (in IEEE Trans. on Computers - February 1971 pag. 211), provato da DAVIES (in IEEE Trans. on Computers - February 1972 pag. 187), con qualche leggera modifica dato il diverso dominio di definizione.

Consideriamo innanzi tutto le funzioni di Rademacher

$$R_j(S)$$

cioè le funzioni che assumono valore 1 se il j-esimo bit di S (al solito scritto in binario) è 0, -1 se è 1.

Per esempio, le prime funzioni di Rademacher sono illustrate nella figura di pag. 6.



Se ne comprende quindi facilmente la struttura. Esse costituiscono un insieme ortonormale, ma non completo.

Ricordiamo poi che il codice di Gray è un particolare codice binario in cui, passando dal numero n a quello n+1 cambia un solo bit. Per esempio i primi 16 numeri sono codificati

	b_4	b_3	b_2	b_1
0	0	0	0	0
1	0	0	0	1
2	0	0	1	1
3	0	0	1	0
4	0	1	1	0
5	0	1	1	1
6	0	1	0	1
7	0	1	0	0
8	1	1	0	0
9	1	1	0	1
10	1	1	1	1
11	1	1	1	0
12	1	0	1	0
13	1	0	1	1
14	1	0	0	1
15	1	0	0	0

La i-esima funzione di Walsh $W_i(S)$ è data da

$$W_i(S) = [b_1 * R'_1(S) \oplus b_2 * R'_2(S) \oplus \dots \oplus b_j * R'_j(S)] * 2-1$$

essendo $b_j\ b_{j-1} \dots\dots b_2\ b_1$ la rappresentazione di i nel codice di Gray e posto $R'_i(S) = \frac{R_i(S)+1}{2}$. Poniamo inoltre $W_0(S)=1$. Si vede quindi che le funzioni di Rademacher sono alcune delle funzioni di Walsh e precisamente

$$R_j(S) = W'_{2^j-1}(S)$$

Data la completezza dell'insieme $\{W_i(S)\}$, possiamo parlare di "trasformata di Walsh" della funzione f(S). Essa è la successione di numeri reali k_i per cui

$$f(S) = \sum_{1}^{\infty} k_i * W_i(S)$$

Per ottenere i k_i basta porre

$$k_i = \int_0^1 f(S) * W_i(S)\, d\, S$$

Si può dimostrare la convergenza di k_i a zero per f(S) continue o che hanno discontinuità in punti con ascissa $m * 2^{-n}$ essendo m ed n numeri interi positivi finiti arbitrari.

Osserviamo innanzi tutto che le funzioni di Walsh (W_i con $i \neq 0$) dividono il segmento unitario in due domini, ovviamente non connessi a parte il caso di W_1, di eguale misura ed individuati dal fatto che nel primo W_i vale 1 e nel secondo vale -1. Ciò significa che possiamo scrivere

$$k_i = \int_0^1 f(S) * W_i(S)\, d\, S = \int_{D_1} f(S)\, d\, S - \int_{D_2} f(S)\, d\, S$$

essendo D_1 e D_2 i due domini. Poichè $D_1 + D_2$ è il segmento unitario, abbiamo anche

$$k_i = \int_0^1 f(S)\, d\, S -2* \int_{D_2} f(S)\, d\, S = k_0 -2 * \int_{D_2} f(S)\, d\, S$$

Questa divisione in due termini è particolarmente significativa nel caso delle funzioni di Rademacher $R_j = W_{2^3-1}$. In tal caso infatti i due domini sono composti completamente di sottodomini connessi di misura 2^{-j}. Come si è visto, $R_j(S)$ assume il valore 1 se il j-esimo bit di S è 0 e -1 se esso è 1.

4 - SEQUENZE LINGUISTICHE BINARIE.

In molte delle applicazioni, la catena di Markov di ordine ω non è un buon modello matematico delle sequenze binarie e ciò perchè

1) in effetti vi sono legami tra i bit a qualsiasi distanza, quindi si dovrebbe avere $\omega \to \infty$;

2) in una catena di Markov di ordine ω, a priori, qualsiasi bit degli ω precedenti quello in esame può influenzare quest'ultimo nello stesso modo, a prescindere dalla distanza, mentre in pratica questa influenza diminuisce in genere con la distanza e, al tendere all'infinito di tale distanza, l'influenza in pratica si annulla.

Nella maggior parte dei casi pratici però ha senso studiare la sequenza come se fosse una catena di Markov di ordine ω, con approssimazione crescente al crescere di ω. Chiameremo "sorgente linguistica binaria" (SLB) una sorgente che gode delle su citate caratteristiche e cioè, riassumento,

a) stazionarietà ed ergodicità,
b) dipendenza tra i bit in genere a qualsiasi distanza,
c) dipendenza in genere decrescente con la distanza e $\to 0$ se la distanza $\to \infty$,
d) approssimazione con catene di Markov di ordine ω tanto migliore, quanto maggiore è ω.[1]

E' stato dato il nome "sorgente linguistica binaria" per analogia con i testi letterari, visti come sequenze di caratteri che godono di proprietà analoghe alle a), b), c), d).

Particolarmente utile è l'uso del "segmento delle sequenze". Le proprietà a),...,d) si riflettono sulla funzione p(S) (che, nel caso markoviano di ordine ω, era costante a tratti). Si ha

A) Nella p(S) di una SLB, comunque dato un numero ε, esiste un numero finito o nullo di discontinuità maggiori di ε.

Un modo di vederlo è il seguente: approssimiamo la p(S) di una SLB con la $p_1(S)$ di una catena di Markov di ordine ω tanto grande che non esista un insieme di misura non nulla per cui sia

$$| p(S) - p_1(S)| < \frac{\omega}{2}$$

---

(1) [Nota (1): Questa proprietà, in effetti, contiene le altre].

Ciò, per ipotesi, deve potersi fare. Quindi è evidente che al più ci saranno 2^{ω} discontinuità maggiori di ε.

Definiamo "quasi-continua" una funzione che goda della proprietà A).

Vediamo ora alcune conseguenze della quasi-continuità.

1) Il numero delle discontinuità della p(S) è discreto (costituisce cioè un insieme di misura nulla); infatti è possibile "numerare" le discontinuità ordinandole a partire dalla più grande (il che può farsi sicuramente data appunto la A).

2) Si può approssimare la p(S), con un errore piccolo a piacere, con una funzione avente un numero finito di discontinuità.

Inoltre le discontinuità maggiori sono in genere poste in punti di ascissa $n * 2^{-m}$ (con n ed m interi), in cui m è (relativamente) piccolo.

Dalla A) si deduce la seguente condizione a cui soddisfano le p(S) delle SLB:

B)- Comunque preso una ε, esiste un α, indipendente da ε, tale che

$$| p(S+\varepsilon) - p(S) | < \alpha * \varepsilon$$

per tutti gli S, a parte un insieme di misura nulla.

Vediamo come ciò si riflette sulla trasformata di Walsh della p(S).

Data la particolare struttura delle funzioni di Walsh $W_i(S)$, si ha che la $W_i(S)$ è costituita da segmenti paralleli all'asse delle S, di lunghezza minima 2^{-j}, essendo j il massimo numero intero minore di $\log_2 i$. Quindi dato $\varepsilon < 1$, esiste un indice $\bar{i}$ tale che, per $i \leq \bar{i}$ si abbia

$$W_i(S) = W_i(S+\varepsilon)$$

[purchè S sia il limite sinistro di un qualsiasi intervallo di costanza della $W_i(S)$]. Essendo $\{K_i\}$ la successione costituente la trasformata di Walsh della

$$| p(S+\varepsilon) - p(S)|$$

si vede che i primi $\bar{i}$ elementi sono nulli. Si può allora scrivere

$$\left| \sum_{\bar{i}}^{\infty} k_i * W_i(S) \right| < \alpha * \frac{1}{\bar{i}}$$

Al diminuire di ε, $\bar{i}$ può aumentare e si ha $\bar{i} \simeq \frac{1}{\varepsilon}$. Passando da $\bar{i}$ a $\bar{i}+1$ si ha

$$\left| \sum_{\bar{i}}^{\infty}{}_i\, k_i * W_i(S) \right| - \left| \sum_{\bar{i}+1}^{\infty}{}_i\, k_i * W_i(S) \right| = \left| k_{\bar{i}} * W_{\bar{i}}(S) \right| = \left| k_{\bar{i}} \right| < \alpha * \left(\frac{1}{\bar{i}} - \frac{1}{\bar{i}+1} \right) < \alpha * \frac{1}{\bar{i}^2}$$

cioè

$$\left| k_i \right| < \frac{\alpha}{i^2} \qquad (1)$$

il che significa che per $i \to \infty$ si deve avere $k_i \to 0$ come $\frac{1}{i^2}$(2). Questa considerazione é importantissima: essa fa sì che basti memorizzare i primi n elementi della trasformata di Walsh per ricostruire, con un certo errore η, la p(S); è evidente che η dipende da n e, quanto maggiore è la memoria, tanto minore è l'errore

La convergenza vista(1) della trasformata di Walsh della p(S) di una SLB significa che, in effetti, ogni stato "dipende" dagli altri "vicini"; detto in altre parole, il numero di gradi di libertà della SLB è molto minore che per un'ipotetica catena di Markov di ordine infinito: in tal caso, essendo 2^ω il numero di stati per una catena di ordine ω, per $\omega \to \infty$, avremmo un'infinità continua di stati indipendenti (e quindi di gradi di libertà) il che non ha senso fisico; la p(S) sarebbe ovunque discontinua e il problema non sarebbe affrontabile matematicamente. Nel nostro caso invece, commettendo un errore piccolo a piacere, può ridursi il numero di gradi di libertà ad un numero finito.

Definiamo la "distanza" ε_{jk} tra due punti del segmento delle sequenze (e degli stati) rispettivamente definiti dalle sequenze di bit $b_1^{(j)}\ b_2^{(j)} \ldots$ e $b_1^{(k)}\ b_2^{(k)} \ldots$

$$\varepsilon_{jk} = \sum_{1}^{\infty}{}_i \left(b_i^{(j)} \oplus b_i^{(k)} \right) * 2^{-i}$$

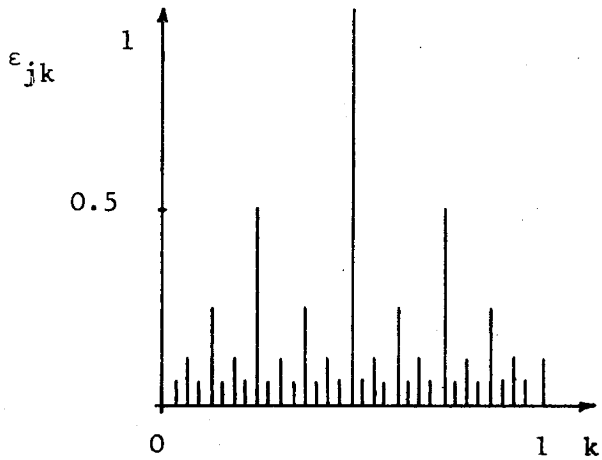
NOTA (1): $\sum_{1}^{\infty}{}_i \frac{1}{i^2} = \frac{\Pi^2}{6}$

NOTA (2): Per una catena di Markov di ordine ω, si ha, per $i \geq 2^\omega$

$$K_i = \int_0^1 p(S) * W_i(S)\, d\, S = 0$$

Si vede che

- $\varepsilon_{jk} = 0$ se, e solo se, $j = k$
- $0 \leq \varepsilon_{jk} \leq 1$
- $\varepsilon_{jk} = 1$ se $j - \frac{1}{2} = \frac{1}{2} - k$, cioè $j + k = 1$
  cioè se j e k sono disposti simmetricamente rispetto al centro del segmento unitario.
- se i primi n bit di j e k sono eguali si ha $\varepsilon_{jk} \leq 2^{-n}$
- ε_{jk}, per $j \to k$, ha la forma



Ciò posto, può estendersi la proprietà A) di quasi-continuità della p(S), dicendo che

C) - Data una SLB esiste un numero reale β tale che, comunque presi due punti j e k sul segmento unitario, si ha l'importantissima proprietà:

$$| p(j) - p(k) | < \beta * \varepsilon_{jk}$$

Si vede facilmente che dalla C) seguono la A) e la B). Data la nostra definizione di ε_{jk}, il viceversa non è generalmente vero (basta pensare ad una discontinuità finita in un punto d'ascissa "irrazionale" o almeno non esprimibile con $m * 2^{-n}$ essendo m ed n finiti).

5 - PROBABILITA' DI PRESENZA P(S).

Data una SLB (e cioè, data la sua p(S)) può definirsi una funzione P(S), tale che, presi sul segmento unitario delle sequenze due punti S_1 ed S_2 con $S_1 \leq S_2$, la probabilità che una sequenza generata dalla SLB data si trovi tra quelle comprese nell'intervallo (S_1, S_2) è

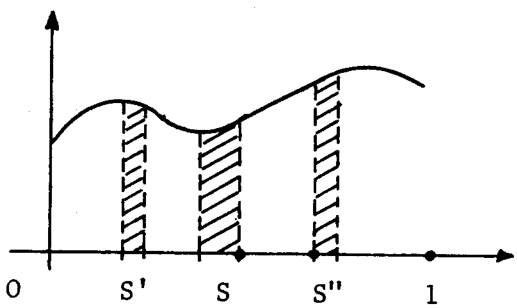
$$\mathrm{Prob}(S_1 \leq S \leq S_2) = \int_{S_1}^{S_2} P(S)\, d\, S$$

Ovviamente deve essere

$$\int_0^1 P(S)\, d\, S = 1$$

Ricaviamo il legame tra p(S) e P(S).

In un passo della sequenza, essendosi nell'intervallo $(S, S+\Delta S)$, si può andare nell'intervallo $(S', S'+\Delta S')$ o nell'intervallo $(S'', S''+\Delta S'')$ a seconda che segua rispettivamente 0 o 1,



essendo

$$S' = \frac{S}{2}$$

$$S'' = \frac{S+1}{2}$$

$$\Delta S' = \Delta S'' = \frac{\Delta S}{2}$$

Si può perciò scrivere, nota p(S),

$$\int_S^{S+\Delta S} p(S) * P(S)\, d\, S = \int_{S''}^{S''+\Delta S''} P(S)\, d\, S$$

$$\int_S^{S+\Delta S} [1-p(S)] * P(S)\, d\, S = \int_{S'}^{S'+\Delta S'} P(S)\, d\, S$$

facendo il limite per $\Delta S \to 0$, per la continuità della p(S), può porsi

$$1) \quad \begin{cases} 2 * p(S) * P(S) = P(\frac{S+1}{2}) \\ \\ 2 * [1-p(S)] * P(S) = P(\frac{S}{2}) \end{cases}$$

Si può dimostrare che questo sistema funzionale non ammette soluzioni P(S) continue o quasi continue, a parte il caso banale di p(S)=1/2. Infatti, da 1) si h

$$P(S) = \frac{P(\frac{S}{2}) + P(\frac{S+1}{2})}{2}$$

Si dimostra, con una dimostrazione più tediosa che interessante, che l'unica funzione continua o con numero discreto di discontinuità che la soddisfa è la costante.

Allo stesso risultato si può arrivare osservando che gli insiemi delle sequenze standard e non standard sono ovunque densi sul segmento unitario(1). Infatt se S è una sequenza di lunghezza infinita standard (o non standard), così pure è S^* che ha i primi n bit diversi da S; quindi comunque preso un punto R sul segment unitario, si può costruire una S^* tale che $|R-S^*|<\varepsilon$, comunque preso $\varepsilon > 0$; ricordi mo infine che sia l'insieme delle sequenze standard, sia quello delle sequenze non standard, hanno un'infinità continua di elementi, per $0 < H < 1$.

La presenza di questa infinità continua di discontinuità pone certe difficoltà al suo studio. Teoricamente gli integrali scritti non avrebbero senso analitico ma ciò non ci interessa per l'uso che ne faremo, dato che ciò che in genere utilizzeremo sarà $Prob(S_1 \leq S \leq S_2)$ che è un numero con un significato fisico definito, che può calcolarsi(2), tra l'altro, col metodo proposto al Paragrafo 1.

---

(1) NOTA (1): Nell'ipotesi di $0 < p(S) < 1$.

(2) NOTA (2): Esattamente per valori di S_1 ed S_2 esprimibili nella forma $m*2^{-n}$ con m ed n interi finiti; approssimativamente negli altri casi.

6 - TRASFORMATE DI WALSH DELLA p(S) E DELLA P(S).

Per quanto detto alla fine del Paragrafo 3, si ha

1) $$\int_0^1 P(S) * R_0(S)\, d\, S = 1$$

2) $$\int_0^1 P(S) * R_i(S)\, d\, S = \text{Prob}(0) - \text{Prob}(1) = \text{costante al variare di } i$$

Anche da ciò si deduce che, se non è

$$\text{Prob}(1) = \text{Prob}(0) = \frac{1}{2}$$

P(S) non può che essere ovunque discontinua: per mostrarlo basta considerare il limite per $i \to \infty$ della 2).

Ricordiamo ora che l'autocorrelazione $A_B(j)$ di una sequenza

$$B = b_1\ b_2 \ldots\ b_i$$

può porsi come

$$A_B(j) = \text{Prob}(b_{i-j}=1,\ b_i=1) + \text{Prob}(b_{i-j}=0,\ b_i=0) -$$
$$- \text{Prob}(b_{i-j}=0,\ b_i=1) - \text{Prob}(b_{i-j}=1,\ b_i=0)$$

Notiamo poi che

$$I_j^1 = \int_{\frac{1}{2}}^{1} P(S) * R_{j+1}(S)\, d\, S = \text{Prob}(b_{i-j}=0,\ b_i=1) - \text{Prob}(b_{i-j}=1, b_i=1)$$

$$I_j^0 = \int_0^{\frac{1}{2}} P(S) * R_{j+1}(S)\, d\, S = \text{Prob}(b_{i-j}=0,\ b_i=0) - \text{Prob}(b_{i-j}=1, b_i=0)$$

e per $j=0$

$$I_0^1 = \int_{\frac{1}{2}}^{1} - P(S) = -\text{Prob}(1)$$

$$I_0^0 = \int_0^{\frac{1}{2}} P(S) = \text{Prob}(0)$$

Si ha, per $A_B(j)$,

$$A_B(j) = I_j^0 - I_j^1$$

usando le relazioni funzionali del Paragrafo 5, si ha

$$I_j^1 = \int_{\frac{1}{2}}^{1} P(S) * R_{j+1}(S)\, d\, S = \int_0^1 P(\tfrac{S+1}{2}) * R_{j+1}(\tfrac{S+1}{2})\, d\, \tfrac{S}{2} =$$

$$= \frac{1}{2} \int_0^1 P(\tfrac{S+1}{2}) * R_j(S)\, d\, S = \int_0^1 P(S)*p(S)*R_j(S) d\, S$$

e dato che

$$I_j^0 + I_j^1 = \mathrm{Prob}(0) - \mathrm{Prob}(1) = \mathrm{cost.}$$

si ha

$$A_B(j) = (I_j^0+I_j^1) - 2\, I_j^1 = [\mathrm{Prob}(0)-\mathrm{Prob}(1)]-2 \int_0^1 P(S)*p(S)*R_j(S)\, d\, S$$

Dimostriamo ora che, nel caso di una catena di Markov di ordine ω, se $j<\omega$, $A_B(j)$ è indipendente da j. Basta a tal fine dimostrare che I_j^1 è indipendente da j per $j>\omega$.

Dividiamo l'intervallo (0,1) in 2^ω intervalli uguali adiacenti inferiormente chiusi e indichiamoli con

$$D_i = (i*2^{-\omega},\ (i+1) * 2^{-\omega})$$

con $0 \leq i \leq 2^\omega - 1$.

In D_i p(S) è costante ed eguale a p_i. Chiamiamo inoltre P_i la probabilità di presenza del gruppo di ω bit che sono eguali per tutti gli S dell'intervallo D_i. Si ha

$$I_j^1 = \Sigma_i\, p_i \int_{D_i} P(S) * R_j(S)\, d\, S = \Sigma_i\, p_i * P_i * \int_0^1 P(S) * R_{j-\omega}(S)\, d\, S =$$

$$= [\mathrm{Prob}(0)-\mathrm{Prob}(1)] * \Sigma_i\, p_i * P_i = [\mathrm{Prob}(0)-\mathrm{Prob}(1)] * \mathrm{Prob}(1)$$

che, come si vede, è indipendente da j.

Sostituendo nella formula generale per $A_B(j)$, tenendo conto che Prob(0)=1 -Prob(1), si ha

$$A_B(j) = [Prob(0)-Prob(1)]-2[Prob(0)-Prob(1)]*Prob(1)=$$

$$= [Prob(0)-Prob(1)]*[1-2\ Prob(1)]=$$

$$= [Prob(0)-Prob(1)]^2$$

che è l'autocorrelazione di una sequenza di bit indipendenti. Nel caso generale delle SLB, si può dimostrare, analogamente, che $A_B(j)$ tende asinteticamente a $[Prob(0)-Prob(1)]^2$ al tendere di j all'infinito. La dimostrazione è analoga e si basa sulla quasi-continuità della p(S).

7 - ERGODICITA' E $p^{(i)}(S)$.

Accenniamo al problema dell'ergodicità di una SLB. Condizione sufficiente di ergodicità per una SLB è che sia,

$$0<p(S)<1 \quad \text{per} \quad 0\leq S<1$$

Condizioni sufficienti meno restrittive sono

$$\begin{cases} \lim_{S\to 0} p(S) \neq 0 \\ \lim_{S\to 0} p(S) \neq 1 \\ \text{non esiste tutto un intervallo di } (0,1) \text{ per cui } p(S)=0 \text{ o } p(S)=1. \end{cases}$$

Possono costruirsi insiemi di condizioni sufficienti ancora meno restrittivi sui valori di p(S), ma piuttosto complicate.

Ricaviamo ora la formula che dà la probabilità che, stando nello stato S, si abbia 1 dopo n passi. Tale probabilità sarà indicata con $p^{n)}(S)$. Procediamo per induzione. Si ha

$$p^{(1)}(S) = p(S)$$

$$p^{(n)}(S) = p(S) * p^{(n-1)}(S'')+[1-p(S)] * p^{(n-1)}(S')$$

essendo

$$S' = \frac{S}{2} \quad e \quad S'' = \frac{S+1}{2}$$

Infatti, supponendo di conoscere $p^{(n-1)}(S)$, abbiamo, dopo un passo, probabilità p(S) di avere 1 (e cioè di andare a S") e probabilità [1-p(S)] di avere 0 (e cioè di andare a S').

Si noti che se la sorgente è ergodica, si deve avere

$$\lim_{i\to\infty} p^{(i)}(S) = \mathrm{Prob}(1)$$

cioè deve essere indipendente da S. Ciò può essere usato per controllare l'ergodicità della sorgente (anche quantitativamente).

8 - CONCLUSIONI.

La precedente trattazione illustra un nuovo approccio allo studio delle sequenze binarie. Sviluppi di questo studio, di particolare interesse pratico, saranno ottenuti particolareggiatamente in una prossima nota.

Saranno analizzate le proprietà di particolari classi di SLB.

Un caso particolarmente interessante per esempio è quello in cui la p(S) ha la trasformata di Walsh tutta nulla a parte i termini relativi alle funzioni di Rademacher. Esse saranno chiamate "linearmente predicibili" (SLBLP). E' questo il caso, come si mostrerà, di molte applicazioni pratiche, tra cui, molto importante, quella alle sequenze binarie generate dalla Δ-modulation di una funzione continua.

Data la relativa facilità di studio delle SLBLP, esse potranno essere usate per costruire sequenze binarie con particolari proprietà statistiche.

B I B L I O G R A F I A

[1] Lackey, Meltzer "A Simplified Definition of Walsh Function" I.E.E.E. Trans. on Computers, February 1971 pag. 211.

[2] Davies "On the Definition and Generation of Walsh Functions" I.E.E.E. Trans. on Computers, February 1972 pag. 187.