Claudio Bonechi (clavicordo)

DENTRO AL CUORE DELLE TLC - 1: SORGENTE INFORMATIVA

25 May 2017

Premessa

Nell'articolo "**Il vero cuore delle telecomunicazioni**" viene riportata la relazione del Prof. Castellani in cui vengono prese in esame **3 possibilità per ridurre il tasso di errore** dovuto al rumore nelle trasmissioni, siano esse analogiche o digitali. Il discorso viene svolto a partire da un insieme finito di simboli diversi che la sorgente di informazione è in grado di produrre e di aggregare per formare messaggi fatti di **k simboli** ciascuno (possiamo pensarli in sequenza, uno dietro l'altro).

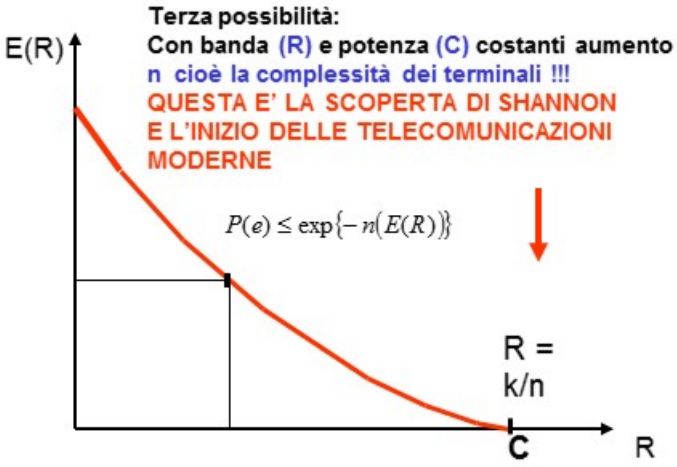
Se l'insieme di simboli è formato da 2 simboli (il numero minimo possibile) che possiamo chiamare "**bit**", il numero dei possibili messaggi sarà 2^k. Quello che interessa alle telecomunicazioni è anche il "ritmo" o velocità di emissione R_e con cui tali simboli vengono trasmessi, cioè
$R_e = k/t \quad \text{bit/s}$.
Per esempio se ogni messaggio viene trasmesso in **1 microsecondo** ed è formato da $k = 8 \text{ bit} \quad R_e = 8/10^{-6} = 8.000.000 \,\text{bit/s} = 8 \,\text{Mbit/s}$ oppure **1.000.000 messaggi/sec**.
Per la trasmissione occorre poi aggiungere ai **k** bit del messaggio originale alcuni bit "di servizio" (secondo i protocolli di comunicazione usati) in modo che i bit diventano in tutto **n** e quindi la velocità di linea diventa $R_e = n/t \quad \text{bit/s}$.

Il rapporto $\boldsymbol{R = k / n}$ sarà quindi sempre < 1 ed è un indice dell'**efficienza della trasmissione**.

Vediamo ora le 3 possibilità:

1. Diminuire il ritmo di trasmissione.
2. Aumentare la potenza del segnale e di conseguenza il rapporto S/N, cioè segnale/rumore.
3. Aumentare la complessità dei terminali

Concentriamoci ora su questa terza possibilità.



Castellani dice:

"Senza cambiare la banda (k / n resta costante) o la potenza del segnale possiamo migliorare le prestazioni facendo **n** sempre più grande. Abbiamo a disposizione una nuova risorsa: la complessità dei terminali. E' evidente che se n è molto grande gli algoritmi di codificazione (e di decodificazione) devono operare su sequenze di dati molto lunghe e quindi sono sempre più complessi.

Come capita spesso per le scoperte scientifiche, la portata rivoluzionaria di questo risultato fu apprezzata solo più tardi, quando la microelettronica ha reso disponibili grandi potenze, grandi velocità di calcolo e grandi capacità di memorizzare i dati.

Lo ha scritto molto bene, nel 1981, un editoriale delle IEEE Transactions on Communications: " Ci troviamo all'alba di una nuova era, quella in cui gli ingegneri possono scambiare banda e potenza per una terza risorsa, la **complessità** di elaborazione. Nel passato, quando i dispositivi attivi costavano cinque dollari l'uno, la parola complessità era gravida di connotazioni negative. Adesso una nuova economia ha capovolto la situazione: l'integrazione a larga scala ha ridotto il costo dei dispositivi di un milione di volte e molteplici nuovi dispositivi sono in attesa dietro le quinte. Pertanto questa **complessità**, opportunamente progettata ed introdotta nei sistemi, diventerà una **necessità economica**".

La sorgente

Per illustrare un po' più ampiamente questa **terza possibilità** occorre prendere in esame il modello generale della comunicazione che prevede una Sorgente di informazione (messaggi), un trasmettitore, un canale di trasmissione, un ricevitore e un ricostruttore dei messaggi della sorgente.

Modello della comunicazione

Sempre pensando che le cose, anche se sono già chiare per tutti , non lo sono mai abbastanza, consideriamo la Sorgente dei messaggi.

Il problema che un sistema di telecomunicazioni deve risolvere è quello di far giungere al destinatario messaggi inviati da un mittente. I messaggi sono aggregazioni di "**simboli di informazione**". Esempio di tali simboli sono le lettere dell'alfabeto, oppure i numeri, oppure elementi geometrici, o suoni: si tratta quindi di "oggetti" che, da soli o in gruppi, rappresentano qualcosa di significativo, qualcosa che costituisce un'**informazione**, o più spesso un pezzo di informazione da unire ad altri pezzi per comporre un'informazione compiuta. Così le lettere dell'alfabeto da sole sono simboli che significano poco, ma raggruppate in certi modi possono significare (ossia portare dei segni) molto. Ma anche un "campione" di segnale (acustico o di altro genere), che poi diventerà un numero binario, è da considerare un "simbolo": da solo non significa nulla ma unito ai campioni precedenti e successivi diventa portatore di significato.

Per poter giungere a destinazione, i simboli devono viaggiare su un **canale fisico** e quindi devono a loro volta essere rappresentati da entità fisiche dette "**segnali**". Un segnale è sempre **qualcosa di fisico** e, direi di più, di analogico, altrimenti non può propagarsi né automaticamente, come nelle telecomunicazioni, né tramite interventi esterni, come l'invio di un messaggero a cavallo.

Ad ogni simbolo (o gruppo di simboli) viene associato un segnale (o gruppo di segnali) tramite un **processo di codifica**, secondo un **codice (ossia un insieme di regole di associazione tra insiemi)**. Il codice deve essere conosciuto sia dal mittente, che nel nostro modello coincide con la sorgente (ma in altri modelli potrebbe anche non coincidere) sia dal destinatario, il quale, tramite il processo inverso di decodifica, potrà ricostruire i messaggi inviati dal mittente.

Ci si rende presto conto che i codici da utilizzare per una trasmissione (e conseguente ricezione) saranno di vario tipo. Possiamo distinguere quindi i codici "fisici", quelli che associano un simbolo

a un segnale (come fa il modem) e/o viceversa (come fa il campionatore), dai codici "logici", quelli che associano un gruppo di bit a un altro gruppo di bit.

In questo articolo ci concentriamo sui simboli e sui codici "logici", che sono oggetto della Teoria dell'Informazione, mentre ci disinteressiamo della teoria dei segnali.

Come si devono trattare i simboli di informazione dal punto di vista delle tecniche di comunicazione e in particolare delle telecomunicazioni? Per inserirli nella progettazione dei sistemi bisogna trovare un criterio quantitativo che valga per tutti i tipi di simbolo, indipendentemente dalla loro natura e dai segnali che li veicoleranno (visivi, acustici, etc.).

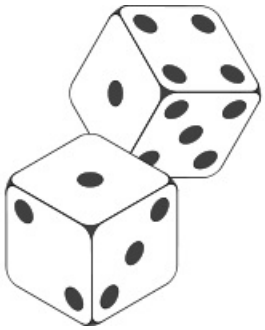
Nella **comunicazione digitale** occorre partire da una sorgente discreta (non nel senso della discrezione, né di quello della qualità…), in grado di emettere un numero finito di simboli, facenti parte di un insieme definito "alfabeto". Se una sorgente è analogica, come il suono e l'immagine visiva, la si può campionare per poi "quantizzare" ogni campione, producendo un numero finito di campioni diversi; quindi, dicendo che ogni campione è un simbolo, non perdiamo di generalità.

Una volta individuati tutti i simboli possibili che una sorgente può emettere, il passo successivo è di **identificare ogni simbolo** (per distinguerlo da tutti gli altri). Tale identificazione consiste nell'**associare un numero ad ogni simbolo** in modo prestabilito, come fa ad esempio il **codice ASCII**. Questa associazione è detta **Codifica di sorgente**. Il numero associato al simbolo è composto da una certa quantità di cifre numeriche: chiaramente, se i simboli emettibili sono ad esempio 3748, servono 4 cifre decimali per esprimere questo numero e quindi per identificare ognuno dei simboli in gioco.

Se invece del sistema decimale si usa il sistema binario le cifre numeriche si chiamano **bit**. Per esprimere in binario il numero 3748 servono 12 bit: 111010100100.

Dunque il numero di bit necessario a specificare i simboli di un certo insieme finito (quello legato alla sorgente) potrebbe essere il criterio quantitativo che cercavamo e, di conseguenza, anche una **misura della quantità di informazione** potenzialmente disponibile alla sorgente. E' quanto abbiamo presupposto più sopra parlando del ritmo di emissione dei simboli.

Misura dell'informazione: poco probabile è interessante



Ma Shannon non si è accontentato di questo ed è andato più in profondità. Ha osservato che la quantità suddetta espressa in numero di bit presuppone che i simboli abbiano tutti la stessa probabilità di essere emessi dalla sorgente, la cosiddetta "**probabilità di occorrenza**". Ma se i simboli hanno probabilità di occorrenza diverse, la quantità di bit necessari a specificare un simbolo può essere, nell'insieme, diversa. Shannon ha quindi proposto l'idea, non nuova, di legare la quantità di informazione "trasportata" da un **simbolo** alla sua **probabilità** di presentarsi.

Parlare di probabilità presuppone parlare di **un insieme di eventi** e non di un evento solo. La probabilità è infatti una descrizione "relativa" e infatti è espressa in % sottintendendo che la somma delle probabilità di un insieme omogeneo di eventi deve essere = 100% cioè =1. Nel nostro caso **l'evento** è l'occorrenza di un simbolo. La probabilità sarà sempre un numero che varia tra 0 (evento impossibile) e 1 (evento certo). Si potrebbe anche dire, invocando un generico principio di dualità, che un insieme di eventi tra loro statisticamente indipendenti è omogeneo se la somma delle loro probabilità di occorrenza è =1.

E' facile convincersi che il concetto di probabilità è del tutto congruo con quello di "informazione". Anche quando noi parliamo, cerchiamo (o dovremmo cercare) di parlare di cose "nuove" ossia poco prevedibili: un messaggio che contiene notizie su **fatti improbabili** è molto **più informativo** di uno che contiene notizie su fatti probabili.

L'analogia è solo formale e non deve confonderci: la teoria dell'informazione non riguarda i contenuti dei messaggi ma solo la loro "espressione", o meglio la ricerca di quella forma dell'espressione che possa garantirne l'integrità durante la trasmissione, oltre a fornire criteri di efficienza nell'impiego dei mezzi di trasmissione. Io penso infatti che il termine "informazione" scelto da Shannon sia un po' infelice, proprio perché rimanda a qualcosa che noi associamo alla conoscenza, ai contenuti dei messaggi.

Tuttavia la scelta di stabilire una sorta di corrispondenza biunivoca tra forma dell'espressione e forma del contenuto (per riprendere la famosa classificazione del semiologo danese Hjelmslev) si è rivelata vincente e ha fatto compiere un enorme salto di qualità alle telecomunicazioni, in cui la forma dell'espressione che consideriamo qui non è altro che la codifica dei messaggi.

Come si configura questa "sorta di corrispondenza biunivoca"? Cercherò ora di spiegarlo.

Autoinformazione

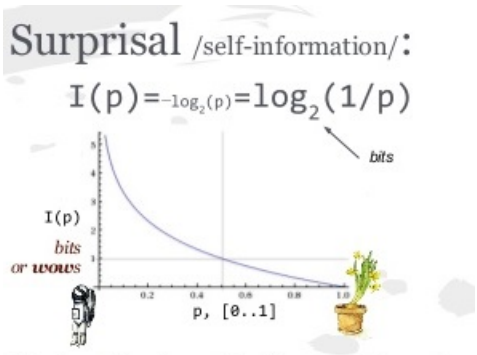
Shannon ha definito la **autoinformazione** (Self Information) **I** associata a un simbolo come una **funzione crescente** al diminuire della sua probabilità di occorrenza p e cioè

$$I = \log_2(1/p) = -\log_2 p \quad \text{bit/simbolo}$$

Questa funzione logaritmica, che era già stata suggerita da Hartley nel 1928, emerge dai requisiti richiesti alla definizione di autoinformazione, il principale dei quali vuole che la autoinformazione associata a eventi **indipendenti** sia la **somma delle autoinformazioni di ciascun evento**:

poiché la probabilità congiunta di due eventi indipendenti è il prodotto delle loro probabilità, ecco che nasce il logaritmo, grande trasformatore di prodotti in somme.

Inoltre essendo la probabilità sempre compresa tra 0 e 1, la **I** sarà sempre **positiva**. Il logaritmo è in base 2 (e d'ora in avanti per tutto l'articolo lo rimarrà) perché così l'informazione viene espressa come successione di alternative binarie rappresentabili in forma di albero gerarchico binario: ad ogni ramo dell'albero è associato un "**bit**" inteso come specificazione di **SI – NO** o anche di **Vero - Falso**. Il bit è quindi un **bit di informazione**. Questo nome fu adottato da Shannon che dice di averlo sentito usare dal matematico John Wilder Tukey come contrazione di "binary **information** digit".



Flipping a fair coin provides 1bit of new information.

Ma chiamare "bit" l'unità di informazione oggi può generare confusione, perché "bit" è entrato nell'uso comune come cifra del sistema numerico a base 2 o binario, anche se questa accezione del termine è venuta **dopo** quella legata all'informazione. Allora, per amore di chiarezza, si è proposto di chiamare il bit numerico "**binit**", nome di origine sconosciuta, almeno a quanto ho potuto sapere io, ma riconducibile a "binary + digit". Di fatto questa distinzione non viene usata nella pratica, se non limitatamente alle trattazioni di teoria dell'informazione. La gente "normale" usa "bit" con significato sia di "cifra numerica" sia di "quantità di informazione", senza rendersi conto che, a rigore, l'equivalenza **quantitativa** esiste solo nel caso di simboli equiprobabili. D'altra parte l'equiprobabilità corrisponde alla massima quantità di informazione (lo si può dimostrare) e quindi, dal punto di vista ingegneristico, quell'equivalenza data per scontata, anche se non genera la massima efficienza, almeno è conservativa, ossia non fa perdere informazione.

Facciamo ora un **esempio di autoinformazione**. Consideriamo un'immagine digitale da **640 x 480 = 307.200 pixel** che contiene **16 livelli di grigio**, ogni livello essendo per noi un simbolo. Se i livelli sono equiprobabili la probabilità di ciascuno di essi in ogni pixel vale **p = 1/16** e l'autoinformazione ***I = log16 = 4 bit/simbolo***; saranno quindi necessari **307.200 x 4 = 1.228.800 bit** di informazione per descrivere la colorazione in grigio dell'immagine. Questo è anche il numero massimo teorico di bit necessario: infatti per le immagini reali non è mai vero che ogni pixel ha una colorazione indipendente da tutti gli altri pixel. Al contrario, qualsiasi immagine è tale proprio perchè ogni pixel è legato a quelli circostanti e i livelli di grigio, cioè i simboli (nel caso presente), non saranno mai equiprobabili.

Entropia della sorgente

Del resto un sistema di comunicazione deve poter accogliere qualsiasi tipo di messaggio, con tutta la variabilità che lo contraddistingue; ciò significa che i simboli di un certo insieme possono venire trasmessi ripetutamente con frequenze statistiche diverse da simbolo a simbolo, o, altrimenti detto, con probabilità di occorrenza distribuite secondo i contenuti dei messaggi, che sono formati da gruppi di simboli o "parole". Anzi, se non ci fosse variabilità nelle frequenze statistiche (che chiamiamo probabilità di occorrenza) non ci sarebbe nemmeno trasmissione di informazione.

Allora quello che serve per progettare un sistema efficiente è la **media dell'informazione**, **pesata secondo le probabilità** di occorrenza di ciascun simbolo, che deve essere trasmessa/ricevuta. Questa media, che la statistica chiama "valore atteso" (expected value) viene detta **Entropia** della sorgente e indicata con **H**: la si misura in **bit/simbolo**. Poiché si parla di un insieme X di M simboli $X = \{x_1, x_2, \ldots, x_M\}$, ognuno dei quali ha una probabilità p_i di essere emesso, si può esplicitare l'entropia

$$H(X) = p_1 \log(1/p_1) + p_2 \log(1/p_2) + \ldots + p_M \log(1/p_M) = -\sum_{i=1}^{M} p_i \log p_i \text{ bit/simbolo.}$$

Se gli M simboli sono equiprobabili

$$p_i = 1/M \rightarrow H(X) = \log M$$

L'insieme **X** costituisce una cosiddetta **"variabile casuale".**

Bisogna precisare l'entropia così definita si riferisce a **simboli statisticamente indipendenti** l'uno dall'altro (cosa che in pratica succede difficilmente). Ciò implica anche che la **sorgente** sia "**senza memoria**". L'altro presupposto fondamentale è che la sorgente sia "**ergodica**", ossia che la sua distribuzione di probabilità non cambi nel tempo: ciò garantisce che la **media di insieme** (con la quale è definita l'entropia) è **uguale alla media nel tempo**. Il nome "entropia", tra l'altro, fu suggerito a Shannon da Von Neumann per l'analogia con l'entropia in termodinamica (da cui differisce solo per il segno -), sostenendo che "tanto nessuno sa bene cosa sia l'entropia".

Nell'esempio dell'immagine di cui sopra, l'Entropia varrebbe **4 bit/simbolo** solo se i simboli fossero **equiprobabili**; in questo caso il numero dei **bit** sarebbe uguale al numero dei **"binit"**. Si può dimostrare che il caso dell'equiprobabilità corrisponde al limite superiore dell'entropia. Per un'immagine "normale" l'entropia si abbassa anche notevolmente dal valore teorico massimo. Ammettendo che l'immagine dell'esempio faccia parte di un video, ciò vuol dire che inviando 25 immagini al secondo, in un secondo invierei circa 1.228.800 x 25 = 31 Mbit/secondo circa per descrivere la colorazione (in grigio) dell'immagine in movimento. Ma se l'Entropia si rivela essere ad esempio 2 bit/simbolo, in un secondo mi basterebbe inviarne la metà, ossia 15,5 Mbit circa (ai quali vanno comunque aggiunti altri bit "di servizio" che descrivono il quadro, etc.). **E' il principio della "compressione" dei messaggi**: significa che posso costruire un codice che mi consente di trasmettere una **quantità di binit** uguale o molto vicina alla **quantità di bit**

calcolata con la formula dell'Entropia, ossia alla forma dei messaggi della sorgente e non solo brutalmente alla loro identificazione "cieca".

Tutto questo corrisponde al buon senso, dal quale discende che si guadagna in **efficienza** se associamo parole di codice più lunghe a simboli meno frequenti (concetto di probabilità). E' quello che ha fatto **Morse** nel 1837 con il suo **codice** per la trasmissione via **telegrafo** . Il codice Morse è una forma di comunicazione digitale ante litteram. Tuttavia, a differenza dei moderni codici binari che usano solo due stati (comunemente rappresentati con 0 e 1), Morse ne usa cinque: punto (•), linea (—), intervallo breve (tra punti e linee all'interno di una lettera), intervallo medio (tra lettere) e intervallo lungo (tra parole). Il criterio di associazione tra lettere – numeri e punti – linee è proprio quello della **frequenza statistica**: alla lettera "e", che in inglese è la più frequente, viene associato un solo punto "•", ossia la parola di codice più breve possibile, mentre al numero "0", viene associata la parola "— — — — —", la più lunga.

Entropy rate

Nelle telecomunicazioni entra in ballo il tempo e quello che serve, più che l'entropia, è la cosiddetta "**entropy rate**" o "information rate" **R = H/t** espressa in **bit/secondo**. Introduciamo anche la **"symbol rate" r** espressa in **simboli/secondo**. Ne segue ***R=r H [simboli/secondo * bit/simbolo = bit/secondo]***

Facciamo un **esempio con il PCM**, a noi più familiare. Consideriamo un segnale x(t) a banda limitata a 50 Hz, campionato alla frequenza di Nyquist cioè a f=100 Hz ossia r = 100 campioni/sec ; ogni campione viene quantizzato in 4 livelli a, b, c, d con probabilità rispettivamente 1/2, 1/4, 1/8, 1/8. Ogni livello può essere visto come un simbolo e il calcolo dell'entropia con la formula scritta sopra dà

$$H = \frac{1}{2}\log 2 + \frac{1}{4}\log 4 + \frac{1}{8}\log 8 + \frac{1}{8}\log 8 = 1{,}75\,\text{bit/simbolo}.$$
$$R = rH = 175\,\text{bit/s}.$$

Ne segue che codificando i 4 livelli con cifre binarie, usando un codice opportuno dovremmo poter trasferire l'informazione media (cioè l'entropia) a 175 binit/s; proviamo allora a usare un semplice codice di sorgente X(a, b, c, d) = 00, 01, 10, 11.

Allora all'uscita del codificatore si avranno dei **binit**, dato che escono numeri , non bit d'informazione. Per conoscere la "binit rate" viene introdotto il concetto che un codice formato da M "parole di codice" ha "lunghezza media" (di parola) $N\_$ così definita:

$$N\_ = \sum_{i=1}^{M} p_i L_i \; con \; L_i = lunghezza \; in \; binit \; della \; parola \; i-esima$$

Nell'esempio sopra $N\_ = \frac{1}{2}\times 2 + \frac{1}{4}\times 2 + \frac{1}{8}\times 2 + \frac{1}{8}\times 2 = 2$ binit/simbolo (in questo caso si vede anche senza calcolo...).

La "binit rate" diventa $r_b = rN_- \quad ossia \quad r_b = 2 \times 100 = 200$ binit/sec.

L'efficienza di questo codice risulta $\frac{H}{N_-} = \frac{1,75}{2} = 0,88$ dunque circa **88%**.
Non è male ma forse si può fare di meglio scegliendo il codice $X(a,b,c,d) = 0,10,110,111$. Risulta

$N_- = \frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{8} \times 3 + \frac{1}{8} \times 3 = 1,75\,\text{binit/simbolo}$ e **l'efficienza = 1,75/1,75 = 100%.** Abbiamo trovato un codice detto "assolutamente ottimo", perché abbiamo associato parole di codice più lunghe a simboli meno probabili: in questo caso è stato semplice ma in generale non lo è. Esistono però procedimenti consolidati per trovare codici efficienti.

Ridondanza

Finora abbiamo supposto che i simboli emessi dalla sorgente fossero statisitcamente indipendenti. Ne segue che se in un dato alfabeto un simbolo **j** è indipendente da un altro simbolo **i** allora la probabilità **p(j/i)** - ossia la probabilità che si presenti il simbolo **j** dopo che si è presentato il simbolo **i** - rimane = **p(j)**, dato che la presenza di un simbolo non condiziona il simbolo successivo: la probabilità "congiunta" di trovare **i** simboli **j** consecutivi è semplicemente il prodotto delle due probabilità individuali, cioè **p(j,i) = p(j)p(i)**.

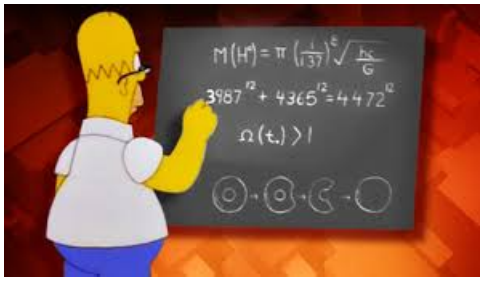
Ma **se l'indipendenza statistica non c'è** allora **p(j,i) = p(j/i)p(i)**, dove **p(j/i) è la probabilità condizionata e si legge "probabilità di j dato i"**.

Per dare un'idea della probabilità condizionata, facciamo due esempi. Primo esempio. In un testo inglese la probabilità di occorrenza della lettera "u", la p(u) è circa 2/100, ma la probabilità che la "u" segua la "q" ossia p(u/q) è quasi =1. Secondo esempio. Si vede bene che la probabilità di estrarre una donna di cuori da un mazzo di 52 carte è 1/52 perché ce n'è una sola nel mazzo, ma ci si può anche chiedere qual è la probabilità che, avendo estratto una carta di cuori, questa sia un donna. La probabilità di estrarre cuori è 13/52 = ¼ mentre la probabilità che avendo estratto cuori la carta sia una donna è 1/13 cioè p(donna/cuori) = 1/13. Si può anche ritrovare la probabilità di estrarre una carta che sia donna e sia di cuori p(donna,cuori) = p(donna/cuori)p(cuori) = 1/13*13/52= 1/52. Si potrebbe anche chiedere qual è la probabilità che, avendo estratto una donna questa sia di cuori. La probabilità di estrarre una donna è 4/52=1/13. La probabilità che sia di cuori è ¼. Di nuovo, p(donna,cuori) = p(cuori,donna) = p(cuori/donna)p(donna)=1/13*1/4 = 1/52.

Nel caso dei messaggi bisognerà considerare la maggior parte delle influenze di un simbolo o un gruppo di simboli sugli altri simboli o altri gruppi di simboli, facendo un'analisi delle probabilità condizionate; considerarle tutte può essere molto pesante e in genere ci si ferma a quelle che danno contributi significativi. L'analisi può anche essere molto complessa. La stessa formula dell'entropia H vista prima, se costruita con le probabilità condizionate, conduce **all'entropia condizionata Hc**; quest'ultima, dato che le probabilità condizionate sono maggiori di quelle non condizionate, sarà sempre minore di quella non condizionata: **Hc <H.**

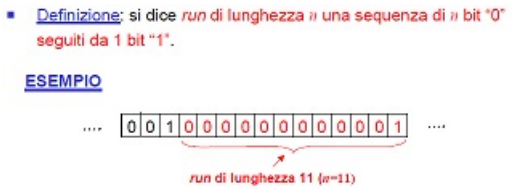
Se i simboli di un messaggio sono statisticamente indipendenti **Hc = 0**. Se invece Hc >0 è segno che nel messaggio c'è della **ridondanza**, ossia c'è informazione non necessaria alla comprensione del messaggio stesso. Per altro verso la ridondanza è un modo efficace di **proteggere il messaggio** perché in caso di alterazione di alcuni simboli il messaggio è ancora decodificabile (comprensibile).

Io pos o cap re una fr se an he se tolg alc ne le tere p rché a li gua itali na è rid ndante. Come è noto la natura usa molta ridondanza negli organismi biologici e d'altra parte non ha problemi di efficienza. Non è così per noi umani, perché **la memoria consuma spazio e tempo e quindi costa**; allora sono stati studiati vari metodi per ridurre la ridondanza, in particolare nell'informatica, dove sono saltati fuori i formati zip, rar, mp3, mp4, jpeg, flac, etc. Certo, man mano che la memoria costa di meno e gli elaboratori e i canali di trasmissione diventano più veloci, questo risparmio perde un po' di appeal. Un po' come succedeva (allarme amarcord!) con i programmi di computer in tempi passati, quando per risparmiare memoria erano così sintetici da risultare incomprensibili per uno che tentava di capirli.



Un insieme di simboli e di messaggi privo quasi del tutto di ridondanza è la **matematica**: in generale non è possibile eliminare nemmeno un simbolo da un'espressione matematica senza stravolgerla. Ma in generale i messaggi sono quasi sempre ridondanti, a vario grado.

Un esempio di codifica di sorgente che **riduce la ridondanza** è la **"codifica run-length"** usata quando il messaggio binario contiene lunghe sequenze di "zeri" o di "uni".



Invece di trasmettere l'intera stringa, si trasmette il valore n che caratterizza la sua lunghezza: usando parole di codice composte da k bit, è possibile rappresentare stringhe con lunghezza massima pari a $n = 2^k - 1$ diminuendo drasticamente la ridondanza e quindi impiegando meno binit.

La ridondanza è definita come $\gamma = 1 - \frac{H(X)}{N\_}$

Codifica di sorgente discreta

Con il **Primo Teorema di Shannon** si può dimostrare che per una sorgente discreta senza memoria la relazione tra entropia H(X) e lunghezza media delle parole di codice $N\_$ è:

$N\_ \geq H(X)$
e che anzi si ha
$H(X) \leq N\_ \leq H(X) + 1$
Un **codice** per cui vale $N\_ = H(X)$ si dice **assolutamente ottimo**.

Un codice utilizzabile non deve generare ambiguità in sede di decodifica: in tal caso è detto **univocamente decifrabile (ud)**.
Esempio 1. Il codice $A = 0,\ B = 1,\ C = 10,\ D = 11$ non è **ud**, perché la sequenza 10011 può essere decodificata come *BAABB* o *CAD* o *CABB*.

Un codice ud si dice **istantaneo** o **"a prefisso"** se ogni parola di codice è scelta in modo tale che **non risulti prefisso** di altre parole di codice. Ciò è realizzato nell'esempio 2.

Esempio 2. $A = 0,\ B = 10,\ C = 110,\ D = 111.$

Nell'Esempio 1 la B è prefisso sia di C che di D, mentre nell'Esempio 2 ciò non accade per nessun simbolo. Il non essere prefisso è necessario perchè, per motivi di efficienza, le sorgenti digitali in generale non fanno uso di separatori tra parole di codice. Le lingue scritte occidentali fanno uso del separatore di parole "spazio": esso non era usato in latino fino al IX secolo, rendendo più difficile la decodifica e quindi la comprensione di un testo scritto. Le lingue orientali come il cinese, il giapponese e il coreano non fanno uso di spazi. La nostra lingua parlata usa le pause come separatori tra frasi o parti di esse. Dato che molte parole sono prefisso di altre, la comprensione prevede la conoscenza del contesto, ossia la conoscenza, oltre che del vocabolario, anche della grammatica e della semantica.

A seconda che la sorgente abbia o non abbia memoria, la strategia di codifica segue strade diverse. Se ha memoria si cerca prima di togliere la ridondanza (compressione senza perdita, come il formato "rar") per trasformarla in sorgente senza memoria. Dopodiché la codifica assegnerà parole di codice più lunghe ai simboli meno probabili, cercando di raggiungere il miglior compromesso tra efficienza del codice, ritardo di codifica, complessità di elaborazione, efficienza nella decodifica (ricostruzione del messaggio).

Uno dei codici sorgente più usati è il **Codice di Huffmann** che è un **codice ottimo,** ossia permette di ottenere il minimo $N\_$ per una data sorgente, quello che più si avvicina all'entropia.

Trasmissione dei messaggi

Poiché un messaggio deve essere trasmesso attraverso un canale, sarà necessario usare un codice che renda il messaggio il più possibile adatto a quel canale. Sarà cioè necessario usare una codifica di canale. Ma prima di inviare un messaggio al codificatore di canale è buona norma "ripulirlo dalla ridondanza", codificando la sorgente per minimizzarne la ridondanza e ottenere la massime efficienza possibile.

Un buon sistema di trasmissione richiede comunque una sorgente il più possibile efficiente non tanto per la sorgente in sé quanto per l'ottimizzazione della trasmissione dei simboli su un canale, secondo la "terza via" illustrata da Castellani. Questa ottimizzazione richiede due step: 1) massimizzazione dell'entropia della sorgente togliendone la ridondanza, tramite la codifica di sorgente; 2) introduzione "controllata" della ridondanza tramite la codifica di canale.

La teoria dei codici di sorgente ha diversi aspetti, alcuni dei quali sono stati qui solo sfiorati, altri sono stati ignorati. Ma era tanto per dare un'idea sia pure vaga della problematica. Essa costituisce un **componente della complessità** degli apparati, concetto richiamato nella citazione IEEE riportata all'inizio. Torneremo sulla **complessità** come approccio al problema della comunicazione digitale nel prossimo articolo, in cui affronteremo la **codifica di canale**.

Estratto da "http://www.electroyou.it/mediawiki/index.php?title=UsersPages:Clavicordo:dentro-al-cuore-delle-tlc"