



Claudio Bonechi (clavicordo)

TEORIA DELL'INFORMAZIONE E SIGNIFICATO

14 March 2011

Premessa

Dal punto di vista puramente tecnico, la grandezza "Informazione" (così come usata nella Teoria dell'Informazione di Shannon) ha poco a che vedere con il concetto di informazione comunemente inteso, che si riferisce alle notizie e al sapere in generale. Ciò ha generato una certa ambiguità nell'uso di questo termine, generando equivoci o fraintendimenti.

Informazione tecnica ed informazione comune

Per chiarezza, chiamiamo **"informazione tecnica"** la grandezza "(quantità di) informazione" elaborata da Shannon, chiamata generalmente "I" e misurata in bit. Chiamiamo **"informazione comune"** l'informazione come viene intesa nei media e nella vita quotidiana: in sé non è una grandezza fisica, ma, per noi esseri umani, qualcosa di mentale, che in generale ha (o avrebbe) lo scopo di aumentare la nostra conoscenza.

L'uomo, è stato detto, è per natura un "essere sociale" e questo implica tra l'altro che noi tutti abbiamo l'esigenza di "comunicare", ossia di condividere, mettere in comune qualcosa di interno (mentale, biologico, affettivo) con il mondo esterno, in particolare quello umano.

Ci sono vari modi in cui la comunicazione interumana può avvenire: variando il tono della voce, toccando una persona, parlando, muovendosi, scrivendo, suonando, disegnando

Ciascuno di questi modi prevede l'emissione e la ricezione di dati, che possono essere immediati, come in una radiocronaca, oppure memorizzati in precedenza. L'"informazione comune" è il risultato mentale ottenuto dalla comunicazione di un insieme "strutturato" di dati, ai quali un gruppo sociale o un individuo attribuiscono un valore. Tipico esempio di dati che vengono strutturati sono le parole del linguaggio verbale.

Dato, struttura, significato

Un "dato" (dal latino "datum" che significa letteralmente "fatto", "accadimento") è una rappresentazione elementare, spesso codificata, di una cosa, di una transazione,

di un evento o di altro considerato rilevante per la conoscenza. Il dato, nella nostra accezione, non è l'oggetto, ma una sua rappresentazione simbolica. Un insieme strutturato è un insieme i cui elementi sono caratterizzati da legami che rispondono a regole sintattiche, ovvero a regole di associazione interna, implicite o esplicite (o un misto delle due).

Un insieme strutturato di dati può perseguire una finalità semantica, ossia la rappresentazione di uno o più significati. I dati assumono allora il ruolo di "segni", e possono diventare mezzi per dare vita a una comunicazione. La comunicazione è uno scambio di segni secondo un insieme di regole, detto "codice". Sappiamo che un segno è costituito da due parti: un "significante", elemento formale basato su un supporto fisico, e un "significato", elemento concettuale che rimanda a un oggetto esterno; i due elementi sono correlati e la loro correlazione è stata attuata da un qualche contesto sociale: le regole di questa correlazione fanno anch'esse parte del codice. Il significante può essere pensato come contenente e il significato come contenuto.

Un significato emerge (ossia dà luogo a un segno) perché risponde a certe esigenze di valore per un certo gruppo sociale: diventa un possibile costituente dell'informazione comune, la quale risulta da un insieme di segni legati tra loro a comporre "messaggi", disposti spesso in una sequenza, come nella scrittura. Il concetto di informazione umana è stato poi esteso ad altre entità, ad altri esseri viventi e anche a macchine; è stato seguito cioè un modello interpretativo, rivelatosi utile in molti casi, secondo il quale quelle entità si scambiano messaggi (costituiti da raggruppamenti di dati per esse significativi) che hanno lo scopo di attivare processi (costituiti da azioni, comportamenti, cambiamenti di stato, etc.). In questo concetto esteso di informazione comune, chi svolge il ruolo centrale di promotore di tutto ciò che serve a generare e a fruire l'informazione stessa è ancora il significato, o meglio il suo valore (per le macchine da intendere in termini di processo).

La definizione di informazione di Shannon

Discorso diverso è quello che riguarda la "informazione tecnica", quella definita da Shannon, misurata in bit. Essa non ha assolutamente nulla a che fare con il significato ma si applica solo ed esclusivamente al significante. Quest'ultimo è l'elemento formale che necessita di un supporto fisico, sia per la memorizzazione che per la comunicazione della "informazione comune". E' nell'ambito del significante e delle sue caratteristiche formali e fisiche che si svolge il discorso (e la teoria) sulla informazione tecnica. Il supporto fisico di un significante si compone di due parti: il canale e il segnale. Il canale è tipicamente qualcosa di statico che supporta il transito (trasmissione) o la permanenza (memorizzazione) del segnale. Possono svolgere il ruolo di canale: lo spazio vuoto senza confini, una porzione di spazio (vuoto o pieno di aria o di altra materia), un cavo metallico, una fibra ottica, un foglio di carta, un

tubo pieno di fluido, un hard disk, etc. Il segnale un elemento di energia. Può essere statico, come i grafismi utilizzati nella scrittura tradizionale (lettere dell'alfabeto, numeri e altro), la magnetizzazione distribuita su supporti magnetici, le variazioni di superficie o di volume (incisioni o rilievi) su supporti di vario tipo; nel segnale statico l'energia in gioco è utilizzata "una tantum" per la "scrittura" (le virgolette indicano che va intesa in senso generalizzato), mentre per la lettura è utilizzata ogni volta che serve. Il segnale può essere dinamico se utilizza energia che si propaga; la propagazione può avvenire per onde, elettromagnetiche o acustiche, o in altri modi, meccanici e chimici. In ogni caso, al concetto di segnale è sempre associato il concetto di variazione. Il fatto che l'energia associata a un segnale possa assumere forme diverse, rende possibile assegnare al segnale il ruolo di significante.

Le forme devono essere ben distinte affinché siano il meno possibile ambigue, il meno possibile fonte di errori di individuazione. Per spostare l'accento più sulla comunicazione che sulla tecnica, è meglio chiamare "simboli" i segnali, dando per scontato che ad essi verranno associati dei significati; infatti nella disciplina che studia i segni (la semiotica) il simbolo (distinto dall'immagine e dall'icona secondo Pierce) è un tipo di segno. Nella pratica i simboli statici hanno soprattutto il compito di memorizzare i dati. I simboli dinamici hanno soprattutto il compito di trasportare i dati e corrispondono quindi a segnali acustici o elettromagnetici, fatti di onde di energia in grado di propagarsi in vario modo nello spazio. Nel supporto fisico, relativamente alla informazione tecnica, occorre considerare due aspetti rilevanti: l'efficienza e l'efficacia.

L'efficienza è connessa all'energia da mettere in gioco. E' stato dimostrato che non è possibile memorizzare o comunicare informazione senza impiego di energia; ovviamente è opportuno ridurre il più possibile la quantità. In generale per un singolo messaggio, un messaggero a cavallo consuma molta più energia di un sistema di trasmissione radio. Ma l'efficienza riguarda anche la quantità di simboli necessari a memorizzare o trasmettere/ricevere i dati, e ciò non solo per l'energia coinvolta (più simboli, più energia) ma anche per l'economia complessiva del sistema di memorizzazione-ricetrasmisione, sia nei supporti di memoria sia nel canale di ricetrasmisione. Ad esempio la banda di frequenza radio è un bene comune e va sfruttata al meglio. Poiché la tecnica digitale è basata sull'associazione tra simboli e cifre numeriche (in particolare binarie, ma concettualmente questo è irrilevante) diventa importante chiedersi quanti simboli sono necessari per descrivere (e quindi memorizzare e trasmettere/ricevere) una certa quantità di dati e quante cifre numeriche binarie (bit) sono da associare quei simboli.

La risposta a questa domanda è cruciale per i sistemi di comunicazione e di trattamento dei dati; per formularla Shannon (ma anche Wiener, "padre" della cibernetica, partendo da altri presupposti è andato nella stessa direzione) ha introdotto la statistica nella teoria della comunicazione e ha costruito la Teoria dell'Informazione: fatto che ha portato conseguenze epocali nelle comunicazioni elettriche. Ma come effetto collaterale il nome "Informazione" è diventato una fonte

di ambiguità a danno dei non specialisti (certamente non della "casalinga di Voghera"!).

Simboli ed alfabeto

Il punto di partenza è semplice: dato che si presuppone che qualsiasi fonte di informazione sia descrivibile tramite simboli, per farlo occorre un "alfabeto" di simboli, ossia una collezione di simboli tutti diversi tra loro. Ad esempio, se voglio descrivere un disegno in bianco, grigio e nero, posso suddividere il disegno in una griglia di quadratini, facendo diventare ogni quadratino un simbolo. Si possono numerare i quadratini e individuare in ogni quadratino un livello di grigio da 0 (bianco) a 9 (nero) con i numeri da 1 a 8 che indicano un grigio sempre più scuro. Sto usando implicitamente il sistema decimale, quello imparato a scuola da bambino, che usa 10 cifre (da 0 a 9) ma potrei usare il sistema binario come fanno i computer. In questo esempio mi servono due gruppi di cifre numeriche, uno che indica la posizione del quadratino e l'altro che indica il suo livello di grigio; per il primo gruppo, se ho scelto una griglia di 1000 quadratini, servono 3 cifre (da 000 a 999), mentre per il secondo ne basta una che indica il livello di grigio; in tutto necessito di 4 cifre per ogni quadratino, cioè 4 cifre/simbolo. Sappiamo che $3 = \log_{10}(1.000)$; più in generale, detto N il numero dei simboli (quadrantini, o livelli sonori, o elementi di un qualunque insieme a cui è stato assegnato il ruolo di simbolo), diciamo che $\log_{10}N$ è il numero di cifre decimali necessario per individuarli tutti, mentre $\log_2N$ è il corrispondente numero di bit, o "bit/simbolo". Entrambi i numeri sono da approssimare per eccesso all'intero, dato che il numero di cifre non può che essere intero.

Fin qui niente di strano: operando come nell'esempio dei quadratini, l'associazione tra numero e simbolo è molto semplice. Tale associazione tra simbolo e numero viene detta comunemente "codice", che in questo caso è addirittura banale. Nel mondo della comunicazione digitale il procedimento di codifica consiste nello stabilire delle regole di corrispondenza tra due insiemi, quello dei simboli e quello dei numeri, regole che devono stabilire in che modo assegnare i numeri ai simboli. Questa assegnazione potrebbe essere una corrispondenza uno a uno, ma potrebbe essere anche diversa, per esempio un numero potrebbe rappresentare non un simbolo ma un gruppo di simboli.

Naturalmente a monte di tutto questo ci sarà stato un procedimento che ha creato un "alfabeto" di un numero finito simboli che rappresentino adeguatamente la realtà di interesse: quadratini (o rettangolini) per un disegno, un'immagine (statica o in movimento), livelli sonori per la voce e la musica, livelli di altre grandezze fisiche quali la temperatura, la pressione, etc. Una volta stabilito il codice, ogni fonte di informazione comune diventa una fonte di numeri, preferibilmente binari, ossia bit. Non è detto che la corrispondenza tra simboli e numeri debba essere uno a uno. L'esempio più classico è forse quello del codice Morse, il quale nell'assegnare le

combinazioni di punti e linee (che svolgono lo stesso ruolo delle cifre binarie) alle lettere dell'alfabeto, cercò di massimizzare l'efficienza: basandosi su uno studio statistico della lingua inglese, assegnò le combinazioni più brevi alle lettere più frequenti. In questo modo, in media, la quantità di punti-linee per simbolo (che in questo caso è una lettera dell'alfabeto) viene minimizzata e con essa gli impulsi del telegrafo. E' in questo senso che Shannon ha introdotto la misura della "quantità di informazione" associata a un singolo simbolo, basandola sulla probabilità di occorrenza del simbolo stesso nel contesto in cui opera: più il simbolo è frequente, cioè maggiore è la sua probabilità di occorrenza, minore è la quantità di informazione "I" ad esso associata, ossia necessaria per individuarlo.

Quantità di informazione

Ma questo non c'entra niente con l'informazione come "notizie, nozioni e simili", che abbiamo chiamato "informazione comune", che interessa noi umani. La "quantità di informazione" I non misura quante notizie noi riceviamo o quanto sono veramente informative le nozioni memorizzate in un libro. Non si interessa dei contenuti, dei significati, del loro valore della informazione comune. Si interessa esclusivamente del supporto fisico all'informazione comune: vuole stabilire la quantità di segnali necessari a veicolare, o a memorizzare, un messaggio informativo quando il campionario di questi segnali sia ridotto solo a due, come con l'alfabeto Morse, fatto di soli punti e linee. Ogni lettera dell'alfabeto, ogni cifra numerica da 0 a 9 viene codificata in una sequenza di punti e linee e così i messaggi diventano sequenze di segnali "binari". Viene considerata l'economia dei segnali, relativamente alla loro quantità. Averla chiamata "quantità di informazione" è stata forse una buona operazione marketing, per così dire, ma ha generato vari equivoci. Per dare un'idea di come la statistica è stata introdotta (Morse l'aveva fatto in modo intuitivo) riprendiamo l'esempio dei quadratini. Se io dico che per specificare fino a 1.000 quadratini servono 3 cifre decimali (o 10 bit, visto che $2^{10} = 1024$) più un'altra cifra per il grigio (altri 4 bit, visto che sono 10 livelli). Immaginando di fare una scansione regolare, per righe o per colonne, la numerazione dei quadratini è del tutto prevedibile; quindi qui applicare la statistica è banale. Ma non è banale applicarla ai livelli di grigio. Se non conosco per niente il disegno, devo assumere che ogni quadratino possa avere un livello di grigio qualsiasi, con probabilità uguale per tutti i livelli; quindi la probabilità associata a ognuno dei livelli di grigio sarà $p=1/10$.

La "quantità di informazione" I associata a un singolo evento è definita come

$$I = \log_2 \frac{1}{p}$$

In questo caso è $I = \log_2 (10) = 3,32$ bit che è la "quantità di informazione" associata a un singolo quadratino.

Ciò significa che per specificare, descrivere, eventualmente riprodurre un quadratino grigio, il cui livello di grigio può essere scelto tra 10 livelli, devo impiegare 3,32 bit. Ora è chiaro che non ha senso parlare di “frazione di bit”: o uso un bit, o ne uso due, o ne uso tre, o ne uso quattro

Ma se penso a 100 quadratini, per specificarli tutti, mi occorrono non 400 bit (4 bit per quadratino come detto poco fa), ma 332, con un risparmio teorico massimo del 17%. Naturalmente ciò implica l’uso di una codifica “a blocchi”, ossia non quadratino per quadratino, ma per gruppi di quadratini, cosa che funziona meglio se i quadratini sono molti: ha senso parlare allora di “quantità di informazione media per simbolo”, che viene detta anche “entropia”. Essa viene calcolata come media delle informazioni singole, tenendo conto delle distribuzioni di probabilità in gioco. Si può dimostrare che l’entropia è massima quando tutti i simboli sono equiprobabili o, detto in altri termini, quando il “disordine” è massimo. In effetti la definizione (che Von Neumann suggerì a Shannon) è uguale a quella dell’entropia termodinamica statistica, con il segno opposto.

Tornando all’esempio, nella pratica non si arriverà a 3,32 bit/simbolo ma con codifiche appropriate certo si potrà restare sotto i 4. Dobbiamo poi considerare che i livelli equiprobabili corrispondono a un’assenza di figura, a una porzione di piano “brizzolata”; se c’è una figura, immediatamente certi livelli di grigio diventano assai più probabili di altri: significa che le probabilità dei livelli di grigio associate a molti quadratini aumentano e quindi diminuisce la quantità di informazione I necessaria per individuare il livello di grigio; ne segue che la quantità di informazione media o entropia diminuisce e potrebbe scendere sotto i 2 bit/simbolo.

Si è detto che misurare la quantità di informazione I in bit corrisponde a suddividere un evento in una successione di alternative binarie; ad esempio se la quantità di informazione necessaria per identificare un evento è di 2 bit, significa che l’evento ha 4 possibili stati, individuabili con le combinazioni 00 01 11 10. Bisogna però aggiungere che si sottintende che i 4 stati siano equiprobabili. Altrimenti la I diminuisce. La misura in bit introduce però un elemento di possibile confusione: il bit non sarebbe un’unità di misura ma una cifra binaria ...

In ogni caso, quando i simboli usati sono più di quelli strettamente necessari, si ha “ridondanza”: il che corrisponde uno spreco, si dirà. Sì, ma è uno spreco che serve a proteggere una sequenza di simboli dalle alterazioni cui va soggetta, prodotte ad esempio dal rumore, nel caso delle trasmissioni, o da deterioramenti elettromagnetici o chimici nel caso della memorizzazione: se qualche simbolo si altera o va perduto, è possibile entro certi limiti ricostruirlo solo se esso era ridondante (su questa possibilità si basa l’arte del restauro).

Quindi la ridondanza, impiegata a profusione dalla natura, ha una funzione protettiva fondamentale, anche se si scontra con l’efficienza. La genialità di Shannon ha dimostrato per via matematica che, a patto di non superare la cosiddetta “capacità del canale” (qualcosa di analogo alla portata massima di un condotto per fluidi), è possibile ricostruire il messaggio originale se esso era stato adeguatamente codificato e se è stato alterato, non importa quanto pesantemente, dal rumore

(per esempio nelle trasmissioni spaziali): un risultato per niente intuitivo. Ciò è ottenuto eliminando dai messaggi in trasmissione la ridondanza (attuando una "compressione") per poi reintrodurla in modo controllato tramite apposite codifiche (dette "di canale"). Dalla trasmissione è possibile estendere questa tecnica anche alla memorizzazione. Quanto detto ha a che vedere con l'efficacia di un sistema di memorizzazione o trasmissione, ossia quella qualità che deve garantire la funzionalità del sistema, i suoi obiettivi.

Efficienza ed efficacia si contrappongono: l'aumento dell'una tende a provocare la diminuzione dell'altra e viceversa. Shannon ha indicato la strada per ottimizzare il rapporto tra le due, e, fatto assai rilevante, ha indicato come superare l'ostacolo del rumore, considerato insormontabile prima di lui (a meno che non fosse contenuto entro certi limiti). Quello che si paga è la necessità di codificare i simboli a blocchi, ciò che nella trasmissione introduce un certo ritardo temporale, perché il codificatore impiega tanto più tempo quanto maggiore è l'ampiezza del blocco da esaminare.

Tornando ai due tipi di informazione, quella comune e quella tecnica, potremmo anche dire che mentre la prima si riferisce al significato, la seconda si riferisce al significante, alla sua integrità, alla sua capacità di trasmettere e memorizzare dei significati, prescindendo dalla loro natura.

Nella esperienza umana reale questo non avviene: significato e significante sono tra loro strettamente legati. Una stessa notizia trasmessa dalla radio o letta in un giornale non ha in noi proprio lo stesso effetto. Ma sul piano tecnico, la possibilità di unificare il trattamento dell'informazione utilizzando la tecnologia digitale ha decisamente ampliato le possibilità comunicative della società.

Estratto da "<http://www.electroyou.it/mediawiki/index.php?title=UsersPages:Clavicordo:teoria-dell-informazione-e-significato>"