

Che cos'è l'informazione?

La parola informazione deriva dal latino in-formare, cioè dare forma a qualcosa. È una di quelle parole, come "tempo" o "energia", di cui pensiamo di conoscere il significato fino a quando non ci viene chiesto di definirle con precisione¹. Sta di fatto che il termine "informazione" ha moltissime accezioni, che variano a seconda del contesto in cui vengono usate. Può indicare dati, fatti, notizie, istruzioni, conoscenza, intelligenza, rilevanza, significato, ciò che è rappresentato da una particolare disposizione di segni o simboli, la quantità d'informazione portata da un particolare simbolo, e così via. L'informazione resta uno dei concetti che crea più perplessità anche in chi lavora nel campo. Per il matematico francese René Thom, "il termine 'informazione', troppo carico di tutte le sue connotazioni intenzionali e antropocentriche, dovrebbe essere bandito dalla scienza" (*Dalla morfogenesi alla struttura*). Eppure l'informazione, che non sappiamo ancora definire, può essere elaborata e trasformata da un computer.

Il concetto centrale d'informazione ha a che fare con una particolare relazione tra un "osservatore" e un "evento" che trasmette "informazione" all'osservatore. L'evento è un segno che porta con sé informazione, cioè nuove conoscenze per l'osservatore. Si noti che, se l'osservatore sapesse in anticipo quale evento si verificherà, questo non porterebbe alcuna informazione, ma solo la conferma di quanto già saputo da lui. Se invece l'evento avesse aumentato la sua conoscenza, esso avrebbe trasmesso informazione². Pertanto, l'informazione trasferita da un evento o da un segnale è relativa allo stato di conoscenza dell'osservatore stesso. L'informazione non è qualcosa di fisico che possiamo facilmente misurare, come per esempio la massa di un oggetto. Quindi, è generalmente impossibile parlare in termini assoluti dell'informazione trasmessa da un evento, poiché essa dipende non solo dall'evento, ma anche dall'osservatore. Infatti, lo stesso evento può recare significati diversi a osservatori diversi. Inoltre, anche per lo stesso osservatore, il significato può variare a seconda delle circostanze. Faccio un esempio: supponiamo che mio fratello³ mi venga incontro per strada alzando la mano con il segno del pollice in su e che, oltre a me, altre due persone assistano allo stesso evento. La persona A non sa che cosa significhi il pollice in su, pertanto non ha ricevuto alcuna informazione oggettiva, anche se ha visto quel segno e potrebbe essersi chiesta che cosa volesse dire. La persona B sa che il segno significa "È successo qualcosa di buono", e quindi ha ricevuto dell'informazione oggettiva, ma non sa a chi sia diretto quel segno, né di che cosa si tratti, perciò ha ricevuto solo un po' di informazione. Per me, quel segno significa che l'intervento chirurgico cui si è sottoposta mia cognata è andato bene, e pertanto ne sono felicissimo. Dunque, io ho ricevuto la stessa quantità d'informazione oggettiva di B, e per di più anche una grande quantità di significato soggettivo, in quanto la mia incertezza e il mio stato emotivo ne sono stati drasticamente migliorati. Se supponiamo che lo stesso segno avesse riguardato una scommessa del valore di 10 dollari vinta da mio fratello, il suo impatto per me sarebbe stato molto minore. Questo esempio illustra la natura prettamente soggettiva del significato derivante da eventi oggettivi, ed evidenzia anche due tipi di significato: oggettivo e soggettivo. Il significato oggettivo è il significato formale di un segno, vale a dire quello convenzionalmente condiviso, previo accordo, da una comunità. Il significato soggettivo invece è quello che l'informazione oggettiva ha per un particolare osservatore, ed è

1 Ho più volte paragonato questa nebulosità del concetto di informazione con quanto dice Sant' Agostino del tempo: "Il tempo? So benissimo cosa è ma se tu mi chiedi di spiegartelo io non lo so più". Questa frase è in linea con quanto verrà discusso successivamente riguardo alla difficoltà di comunicare ad altri una nostra conoscenza.

2 Qui va sottolineato che nel processo di trasferimento della informazione ambedue i ruoli comportano una modifica perché si abbia davvero la trasmissione di qualcosa e un uso di energia. A livello macroscopico questo può apparire irrilevante ma non così scendendo via via di dimensioni fino al mondo microscopico.

3 Qui, come accade spesso in semiologia, si fa riferimento implicitamente alla trasmissione di informazione tra due persone. Ma rimane da investigare come la informazione viene inizialmente acquisita dal mondo esterno. In questo caso è difficile ammettere che l'ambiente materiale disponga di informazione che ci trasmette attraverso un evento. In linea con quanto si dice poi è la capacità dell'osservatore che trasforma in informazione ciò che prima era soltanto un accadimento naturale.

condizionato dalla percezione e dal riconoscimento del significato convenzionale del segno. Si noti che la percezione di un segno come qualcosa che “potrebbe avere” significato è indipendente dalla conoscenza del suo significato oggettivo. Questa è la situazione per l'osservatore A che ha percepito il segno: ha capito che forse voleva dire qualcosa, ma non sapeva che cosa significasse. Il riconoscimento del significato oggettivo di un segno può avvenire soltanto se il significato era precedentemente noto grazie a un accordo esplicito o implicito. Pertanto, ciò che chiamiamo informazione si basa sempre su un previo accordo, il che pone dei dubbi sull'universalità di ciò che s'intende per “oggettivo”⁴. È da notare che anche l'“assenza di un evento” può portare informazione e significato. Per esempio, supponiamo che io dica a un amico che, se non mi vede nella piazza principale alle 18, può trovarmi al bar Bellavista. In questo caso sto usando la mia presenza o assenza nello spazio e nel tempo (piazza principale alle 18) come segno per trasmettere una precisa informazione. La mia presenza o assenza corrisponde a un “bit” di informazione avente il significato oggettivo precedentemente concordato: “1” se sono presente, “0” se non lo sono⁵. Presenza e assenza possono quindi trasmettere due significati soggettivi molto diversi tra di loro. Si noti che tutti gli altri osservatori presenti nella piazza principale alle 18 non potrebbero sapere che la mia assenza è informazione per qualcuno! Anche nello spazio vuoto, quindi, potrebbero esserci molte informazioni per gli osservatori che sanno dove e quando osservare.

L'informazione di Shannon

Supponiamo ora di voler esaminare delle informazioni oggettive valide sia per le macchine sia per gli esseri umani. In questo caso i segni o gli eventi che portano informazione hanno significati convenuti, ma non tutti i segni ne convogliano la stessa quantità. Per esempio, un segno di pericolo che appare di rado può contenere molta più informazione rispetto ai segni che si verificano frequentemente. Questo è ciò che Claude Shannon teorizzò nel 1948 con la pubblicazione di un articolo fondamentale intitolato “A Mathematical Theory of Communication”. In esso, Shannon considerò il problema della comunicazione dell'informazione con l'obiettivo di quantificarla.

Egli ne estrasse solo gli aspetti oggettivi e costruì una teoria matematica efficace che è stata fondamentale per ottimizzare l'uso dei canali di comunicazione, per ridurre gli errori e per migliorare l'efficienza dei codici, oltre a consentire molte altre realizzazioni. Il concetto di “quantità d'informazione” si basa sul seguente problema ipotetico: supponi di sapere che qualcuno, un mittente, ti comunicherà una lettera dell'alfabeto inglese. Quando ti arriva, quanta informazione hai ricevuto? Ciò dipende dall'aspettativa che avevi di ricevere quella lettera. Qualora il mittente avesse scelto a caso una carta da un mazzo ben mescolato di ventisei carte in cui ogni carta corrisponde a una lettera dell'alfabeto diversa, ogni lettera avrebbe avuto la stessa probabilità di essere scelta. La probabilità di scegliere un qualsiasi simbolo è quindi uguale a $1/26 = 0,0385$.

Se invece il mittente avesse selezionato a caso una lettera da una pagina di un testo in inglese, la probabilità di scegliere una determinata lettera sarebbe dipesa dalla frequenza con cui quella lettera è usata in inglese. La frequenza relativa delle varie lettere dell'alfabeto inglese varia di molto: dal 12,702% per la lettera “e”, quella più frequente, allo 0,074% per la lettera “z”, la meno usata. Quindi, la lettera “e” ha 172 volte più probabilità di essere scelta rispetto alla lettera “z”⁶. Si noti che, se il testo utilizzato per scegliere a caso le lettere fosse scritto in italiano, la distribuzione delle frequenze relative sarebbe stata molto diversa, anche se l'alfabeto è essenzialmente lo stesso. Perciò la teoria matematica dell'informazione deve presumere che si conosca, o che si possa conoscere, l'alfabeto dei simboli e le loro statistiche. Senza questi chiarimenti preliminari, non è possibile capire il significato ristretto della

4 Per le macchine per cui il discorso è più semplice esiste un alfabeto di segnali fisici che costituiscono il linguaggio e ogni segnale è in corrispondenza biunivoca con una sequenza di bit (parola di codice). Questa è una conoscenza a priori che consente alle macchine in questione di trasferire l'una all'altra una successione di bit che costituisce lo scopo della loro interazione. Per noi è più complesso; la trasmissione attraverso fonemi della parola elettrone, ad esempio, consente il trasferimento di una parola ma essa viene soggettivamente interpretata dai due corrispondenti in modo dipendente dalle loro conoscenze specifiche. Per uno dietro la parola elettrone ci può essere il concetto di funzione d'onda e per l'altro quello di una particella che porta corrente elettrica.

5 E' quanto accade nel RADAR dove l'assenza di un segnale di eco significa che non ci sono aerei nemici.

6 Prima ancora di Shannon queste considerazioni furono usate da Samuel Morse per giungere all'alfabeto che porta il suo nome e che codifica ogni lettera con una sequenza di punti e linee (0 o 1).

parola "informazione" usata nella teoria di Shannon. Si noti anche che conoscere le lettere e la loro probabilità di comparire è molto diverso dal conoscere l'ordine con cui tali lettere sono distribuite nella sequenza. Significa solo che, se si conta il numero di lettere di un certo tipo in un milione di lettere provenienti tutte da un testo inglese, ci si può aspettare di trovare circa 127.020 "e" e 740 "z", senza però conoscerne l'ordine. Nel caso di trasmissione d'informazione a destinatari generici, come in telegrafia o in telefonia, si presume che la trasmissione di ogni lettera sia avvenuta felicemente quando il segnale concordato arriva al destinatario in condizioni tali da essere riconoscibile⁷. Ecco quindi il contesto in cui possiamo calcolare la "quantità d'informazione" portata da un segnale o da un qualsivoglia segno. Shannon definì la quantità d'informazione I_s trasmessa da un segnale s come il negativo del logaritmo in base 2 della probabilità p_s di ricevere quel segnale, cioè: $I_s = -\log_2 p_s$ (il negativo del logaritmo si chiama cologaritmo e si usa spesso quando si fa il logaritmo di un valore compreso tra 0 e 1). Visto che la probabilità è un numero compreso tra 0 e 1, la quantità d'informazione è un numero positivo compreso tra 0 e ∞ . Questa definizione implica che un segnale che ha una piccola probabilità di apparire porta con sé un'elevata quantità d'informazione. Se la probabilità fosse 1, sapremmo esattamente quale segnale comparirà, e in tal caso la quantità di informazioni portata da quel segnale sarebbe 0. Vorrei sottolineare che Shannon non ha mai definito che cosa sia l'informazione, ma ha solo definito la quantità d'informazione portata da un segnale in determinate condizioni. La base 2 del logaritmo è stata scelta in modo che, se l'alfabeto è composto da due soli simboli, per esempio "0" e "1", con uguale probabilità di apparire, la quantità d'informazione portata da ciascun simbolo è 1 "bit" d'informazione (poiché $\log_2 0,5 = -1$). Questa situazione si verifica quando lanciamo una moneta in cui i due possibili risultati hanno la stessa probabilità di uscire. Pertanto, la definizione di cui sopra misura convenientemente la quantità d'informazione in unità di bit.

Non bisogna però confondere il concetto di "bit" come "unità di misura" con il bit utilizzato per indicare il valore di una cifra in un numero binario. Il valore di un bit in un numero binario è solo "0" o "1", mentre la quantità d'informazione portata da un segnale può essere un numero reale, come 2,34 o $\sqrt{2}$. Alcuni studiosi spiegano il concetto d'informazione di Shannon dicendo che essa è proporzionale al grado di sorpresa che suscita nel destinatario quando riceve il messaggio: quanto maggiore è la sorpresa, tanto maggiore è la quantità d'informazione ricevuta. Altre persone, in modo più astratto, affermano che l'informazione è la risoluzione dell'incertezza, il che significa che, quanto maggiore è l'incertezza esistente prima della trasmissione, tanto maggiore è la quantità d'informazione necessaria per risolverla. Ad altri piace descrivere l'informazione come una misura della libertà di scelta a disposizione della persona che seleziona un messaggio. Si noti, però, che la sorpresa, l'incertezza e la libertà di scelta hanno senso solo per enti coscienti e non per una macchina. Una macchina non ha né incertezze né sorprese, e tantomeno libertà di scelta.

Shannon chiamò entropia dell'informazione la media della quantità d'informazione portata da una sequenza di simboli. Sorprendentemente, cambiando il segno della formula dell'entropia dell'informazione si ottiene la formula dell'entropia termodinamica, un concetto fondamentale in fisica. Questa strana "coincidenza" ha portato il concetto di informazione nella fisica⁸, dove continua ad avere un ruolo rilevante. Si noti però che la definizione di Shannon della quantità d'informazione copre solo una piccola parte del concetto d'informazione, perché ne esclude a priori il significato soggettivo, che in fondo è tutto ciò che ci interessa⁹. L'informazione di Shannon equivale ad avere una macchina che misura

⁷ Qui si apre la parte più interessante per i tecnici: quanta energia serve per garantire la riconoscibilità in presenza di disturbi?

⁸ Riguardo ai sostantivi da usare è interessante questa frase attribuita a Shannon: «La mia più grande preoccupazione era come chiamarla. Pensavo di chiamarla *informazione*, ma la parola era fin troppo usata, così decisi di chiamarla *incertezza*. Quando discussi della cosa con John Von Neumann, lui ebbe un'idea migliore. Mi disse che avrei dovuto chiamarla *entropia*, per due motivi: "Innanzitutto, la tua funzione d'incertezza è già nota nella meccanica statistica con quel nome. In secondo luogo, e più significativamente, nessuno sa cosa sia con certezza l'*entropia*, così in una discussione sarai sempre in vantaggio"»

⁹ Si può dire che nella teoria di Shannon non c'è nulla che si riferisca al valore del messaggio per il corrispondente che lo riceve. Questo aspetto è legato alla informazione qui detta soggettiva: l'avviso di una tempesta in un dato punto

i simboli che passano, e dopo aver raccolto le loro statistiche ci dice quanta informazione oggettiva è contenuta in ogni simbolo. Ciò è perfettamente adeguato per una macchina che non può capire il significato soggettivo di alcun simbolo. Una macchina può solo associare una specifica azione preprogrammata a un qualsiasi simbolo o a una qualsiasi sequenza di simboli. Il concetto d'informazione di Shannon richiede solo il corretto rilevamento e riconoscimento di un segno o di un segnale come simbolo di un alfabeto concordato di cui si conosce la statistica. L'informazione è legata soltanto al significato oggettivo di un segno. L'uso della parola simbolo nella teoria dell'informazione dovrebbe implicare un significato, e ciò può far credere che l'informazione di Shannon sia quella che ci interessa, ma non è così. Le parole "segno" o "segnale" sarebbero state più appropriate, poiché la loro connessione con il significato è più debole, sebbene ancora presente. Ma forse "evento" sarebbe stata la parola migliore in assoluto, perché tutto ciò che è richiesto a un osservatore è la percezione e il riconoscimento dell'evento in quanto tale. E "informazione" è allora soltanto il logaritmo della probabilità di osservare quell'evento. Per un essere umano, il riconoscimento di un evento è spesso inconscio ed è solo un prerequisito per avere accesso al ricco significato soggettivo e cosciente associato a esso. Il significato soggettivo di un testo non dipende dalla probabilità d'uso delle lettere, ma dal loro ordine specifico nel formare parole e frasi, e questo significato è esplicitamente escluso nella teoria di Shannon. Nondimeno, questa teoria è stata fondamentale per descrivere alcuni aspetti importanti del comportamento dei sistemi di comunicazione e dei computer.

Che cos'è un bit?

Quasi tutti hanno sentito parlare di bit al giorno d'oggi, anche se per molti il concetto rimane nebuloso. Bit è il nome di un'entità astratta che può manifestarsi soltanto in uno di due possibili stati, come "testa" o "croce" in una moneta. Un bit può quindi essere utilizzato per rappresentare la distinzione più semplice possibile. Per esempio, può indicare sì o no, vero o falso, acceso o spento, sinistra o destra, presente o assente, oppure i numeri "0" e "1" in un sistema numerico in base 2. Ho elencato vari significati quasi-oggettivi che possono essere associati per convenzione ai due stati rappresentati da un bit. Da notare che i numeri che utilizziamo comunemente sono scritti usando un sistema numerico in base 10. Però qualsiasi numero può anche essere scritto in un sistema in base 2 chiamato "sistema numerico binario".

L'informazione di enti coscienti

Le parole non sono identiche alle cose.

Conoscere delle parole relative a dei fatti non equivale in nessun modo alla comprensione diretta e immediata dei fatti stessi. MEISTER ECKHART

L'informazione di Shannon è quella che conta per le macchine inconse, ma non è quella che ci interessa! È fondamentale chiarire questo punto, perché l'ambiguità delle parole usate per descrivere robot e sistemi di intelligenza artificiale (IA) tende a eliminare l'abisso che separa gli esseri umani dalle macchine cosiddette intelligenti. L'informazione che più conta per noi non è quella simbolica, ma è quella semantica, e qui la coscienza è indispensabile. Il riconoscimento dei simboli che, quando si parla di IA, viene fatto passare per comprensione, in realtà è solo una funzione meccanica che facciamo anche noi in automatico. Usare gli stessi termini che si riferiscono alle capacità della coscienza umana per descrivere le macchine è quindi un grave e pericoloso disservizio fatto all'umanità¹⁰.

Quando parliamo, i simboli che adoperiamo servono per trasmettere il significato soggettivo che proviamo, a differenza dei simboli oggettivi usati dalle macchine. Per esempio, per scegliere le parole necessarie per comunicare con Luca, io devo prima convertire il mio significato soggettivo in parole mentali, per poi verbalizzarle. In questo modo creo un'onda sonora che porta il significato veicolato dal linguaggio che entrambi conosciamo. Le onde sonore sono percepite dal sistema uditivo di Luca e riconosciute inconsciamente come le parole-simbolo che intendevo esprimere. Fin qui il processo di riconoscimento è simile a ciò che fanno le macchine. Subito dopo, però, Luca sperimenta nella sua

dell'oceano può essere recepita da tante barche ma per quelle nelle vicinanze del fenomeno il valore è enorme e può significare la differenza tra la vita e la morte.

10 Per questo motivo ho introdotto il concetto di infosfera per descrivere il comportamento delle macchine senza usare concetti tipici dell'uomo come ad esempio l'intelligenza. L'infosfera può anche applicarsi ai viventi ma solo per quanto riguarda la informazione oggettiva.

coscienza, oltre alla sensazione sonora delle parole riconosciute, anche il loro significato soggettivo, che potrebbe essere diverso dal mio. In sintesi, il riconoscimento di una parola come segno avviene in noi automaticamente, come nei robot, ma nella nostra coscienza viene fatto un passo in più, perché il simbolo viene convertito nei qualia sonori e nel significato a loro associato. Quest'ultimo proviene principalmente dalle emozioni e dai pensieri connessi alla memoria a lungo termine delle nostre esperienze di vita. Il significato da comunicare viene poi automaticamente tradotto in onde sonore, attraverso l'elaborazione dei simboli fatta dal cervello, e trasformata in vibrazioni dai muscoli del corpo. In questo modo comunichiamo significato soggettivo alle persone che ci circondano. Da notare che l'elaborazione cosciente del significato è molto diversa dall'elaborazione automatica dei simboli. Pur sapendone poco, sembra che essa sia basata principalmente su operazioni associative e analogiche tra i qualia, insieme con i processi logici e lineari della mente razionale. Per esempio, il ragionamento fatto con carta e penna quando risolviamo un problema matematico è una combinazione di elaborazione conscia e inconscia, cioè semantica e simbolica, che include pensiero reificato in simboli mentali e materiali, procedure automatiche, intuizioni, emozioni e motivazioni, con la supervisione della coscienza. L'elaborazione dell'informazione che noi trattiamo include anche l'elaborazione del significato che è alla base della nostra esperienza vissuta. Questa elaborazione va ben oltre quella simbolica e algoritmica che possono fare i computer, e che si può acquisire mediante l'informazione di Shannon. Si tratta di una distinzione cruciale, che non è generalmente ben compresa perché la scienza e la tecnologia tendono a enfatizzare soltanto l'aspetto simbolico e meccanico della realtà.

Anche quando un computer utilizza le reti neurali artificiali, il suo comportamento è ancora meccanico, e quindi limitato, perché la comprensione e la creatività vanno oltre ciò che può fare una macchina. Il fatto che i computer possono superare di gran lunga le prestazioni umane nelle attività meccaniche non dovrebbe sorprenderci, perché esiste una complementarità fondamentale tra l'intelligenza umana e quella artificiale. In sintesi, i computer sono strutture fisiche a cui abbiamo trasferito una parte della nostra mente sotto forma di programmi, creando un ponte tra la materia e la mente simbolica e semantica. I programmi ci rivelano soltanto l'aspetto simbolico della mente, perché non posseggono il livello semantico e il libero arbitrio che ci distinguono. I computer sono nostre creazioni dotate soltanto della parte algoritmica della nostra essenza. Ecco perché non siamo computer, ma creatori di computer. La coscienza è la nostra vera ricchezza, perché è ciò che ci permette di comprendere. Usando le parole del filosofo Michael Polanyi, noi "... possiamo conoscere più di quanto possiamo dire" (La conoscenza inespressa).