GIORNALE DI FISICA VOL. LXIII, N. 3 Luglio-Settembre 2022
DOI 10.1393/gdf/i2022-10497-9

PER SAPERNE DI PIÙ

Entropia
Entropy

Giuliano Benenti

Center for Nonlinear and Complex Systems, Dipartimento di Scienza e Alta Tecnologia, Università degli Studi dell'Insubria, Como, Italia

Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Milano, Italia

NEST, Istituto Nanoscienze-CNR, Pisa, Italia

Riassunto. Questo lavoro discute varie sfaccettature dell'entropia, partendo dalla prospettiva originale della termodinamica macroscopica classica, dove il termine è stato coniato, in relazione ai concetti di irreversibilità e di degradazione dell'energia. Viene esaminato anche l'approccio microscopico e probabilistico della fisica statistica. Il paradosso del demone di Maxwell e il principio di Landauer evidenziano poi la relazione tra la termodinamica e la teoria dell'informazione, nella quale l'entropia di Shannon è la grandezza chiave.

Abstract. This work discusses several facets of entropy, starting from the original perspective of classical macroscopic thermodynamics, where the term was coined, and in relation to the concepts of irreversibility and energy degradation. The microscopic and probabilistic approach of statistical physics is also examined. Maxwell's demon paradox and Landauer's principle then highlight the relationship between thermodynamics and information theory, where Shannon entropy is the key quantity.

Non vi è dubbio che il concetto di entropia non sia solamente un cardine della fisica e della teoria dell'informazione ma sia anche entrato nel linguaggio comune come sinonimo di disordine. L'esperienza di ogni giorno ci insegna che l'entropia tende spontaneamente ad aumentare, basti pensare ad una casa dove smettiamo di fare la manutenzione, oppure al nostro tavolo da lavoro, o alla camera dei nostri figli. Per esempio se una tazza cade dal tavolo al pavimento si rompe facilmente, ma i cocci non si rimetteranno insieme spontaneamente per ridarci la tazza integra, cioè non passiamo spontaneamente da un sistema disordinato (la tazza rotta) ad uno ordinato (la tazza integra). Questo introduce una freccia del tempo: se vediamo due foto, una con la tazza integra sul tavolo e l'altra con la tazza rotta per terra, ordiniamo cronologicamente la tazza integra prima di quella rotta. L'irreversibiltà, vale a dire

il fatto che i fenomeni che osserviamo abbiano una direzione del tempo ben definita, si manifesta nonostante il fatto che le leggi elementari della dinamica (le equazioni di Newton) non distinguano il passato dal futuro.

L'irreversibilità è strettamente collegata al concetto di entropia, in quanto l'entropia aumenta nel caso di trasformazioni termodinamiche irreversibili. Il principio di accrescimento dell'entropia seleziona, tra tutte le trasformazioni che conservano l'energia, quelle che possono realmente accadere. Tale principio ha conseguenze fondamentali: le trasformazioni che avvengono spontaneamente in natura hanno un verso determinato e le trasformazioni inverse non si realizzano mai spontaneamente. Come la tazzina che si rompe, oppure il passaggio di calore da un corpo caldo ad uno freddo. Il calore non torna spontaneamente dai corpi freddi a quelli caldi: le macchine frigorifere devono essere alimentate per poter funzionare, serve cioè un dispendio di energia. Come conseguenza il calore ceduto spontaneamente ai corpi più freddi è più difficilmente sfruttabile per produrre lavoro, vale a dire che l'energia si degrada. Pensando all'universo come ad un sistema globalmente isolato, la termodinamica ci dice che evolve aumentando l'entropia, tendendo ad uno stato di temperatura uniforme, dove non saranno più possibili processi energetici, inclusa la vita stessa, chiamato quindi stato di morte termica o morte entropica dell'universo.

1. L'entropia termodinamica

Nel contesto della termodinamica, la funzione di stato entropia (dal greco én. "dentro", e tropé, "trasformazione", vale a dire "trasformazione interna") è stata introdotta da Rudolf Clausius nel 1865, per trovare una risposta alla degradazione dell'energia e all'irreversibilità delle trasformazioni. Clausius espresse in forma sintetica le prime due leggi della termodinamica [1]:

> *... possiamo esprimere nel modo seguente le leggi fondamentali dell'universo, che corrispondono ai due teoremi fondamentali della teoria meccanica del calore:*
> 1. *L'energia dell'universo è costante*,
> 2. *L'entropia dell'universo tende ad un massimo.*

Per definire l'entropia termodinamica, si consideri innanzitutto il teorema di Clausius (o disuguaglianza di Clausius), il quale afferma che, se δQ è il calore che fluisce, in una trasformazione infinitesima, da una sorgenta alla temperatura assoluta T (misurata in gradi Kelvin) al sistema, per una trasformazione ciclica vale la relazione

$$\oint \frac{\delta Q}{T} \leq 0, \tag{1}$$

con l'uguaglianza solo nel caso di trasformazione reversibile ($\oint$ rappresenta l'integrale su un percorso di integrazione chiuso, in questo caso su un ciclo termodinamico). Ricordiamo che una trasformazione reversibile è un'idealizzazione e corrisponde al caso

nel quale ad ogni istante sono ben definiti i valori delle variabili che caratterizzano gli stati di equilibrio del sistema (quali temperatura, pressione e volume) e i valori di tali variabili sono infinitamente vicini a quelli che caratterizzano uno stato di equilibrio. Affinché il sistema sia ad ogni istante infinitamente vicino ad uno stato di equilibrio, la trasformazione deve essere quasi-statica, vale a dire infinitamente lenta (così da poter essere vista come una successione di stati di equilibrio). Inoltre in una trasformazione reversibile assumiamo che il verso della trasformazione possa essere invertito cambiando infinitamente poco i fattori che l'hanno determinata. È chiaro che nella realtà esistono solo trasformazioni irreversibili. Ad esempio le forze di attrito impediranno la reversibilità della trasformazione. Il concetto di trasformazione reversibile costituisce comunque un caso limite delle trasformazioni irreversibili e in certi casi (trasformazioni molto lente, con attriti in prima approssimazione trascurabili) una trasformazione reversibile può costituire una buona approssimazione di una trasformazione reale e quindi irreversibile.

Dato che per ogni trasformazione reversibile l'integrale esteso ad un ciclo di $\delta Q/T$ è nullo, possiamo definire una funzione di stato S, chiamata entropia, in modo che la sua variazione tra due stati A e B del sistema

$$S(B) - S(A) = \int_A^B \left(\frac{\delta Q}{T}\right)_{\text{rev}}, \tag{2}$$

dipenda solo dallo stato iniziale A e dallo stato finale B e non dalla particolare trasformazione eseguita per andare da A a B; infine si è scritto "rev" a pedice di $\delta Q/T$, dato che l'equazione scritta sopra ha senso solo per trasformazioni reversibili.

Le quantità δQ e T sono però ben definite anche per trasformazioni irreversibili. Consideriamo quindi un ciclo termodinamico costituito da una trasformazione irreversibile γ da A a B e una reversibile δ da B ad A. Siccome il ciclo così costruito è globalmente irreversibile, dal teorema di Clausius

$$\oint \frac{\delta Q}{T} = \int_{A\gamma B} \left(\frac{\delta Q}{T}\right)_{\text{irr}} + \int_B^A \left(\frac{\delta Q}{T}\right)_{\text{rev}} < 0. \tag{3}$$

Ne consegue che

$$\int_{A\gamma B} \left(\frac{\delta Q}{T}\right)_{\text{irr}} < -\int_B^A \left(\frac{\delta Q}{T}\right)_{\text{rev}} = \int_A^B \left(\frac{\delta Q}{T}\right)_{\text{rev}} = S(B) - S(A), \tag{4}$$

con il primo integrale che dipende dal cammino (trasformazione) γ percorso. Ne consegue che per una trasformazione adiabatica irreversibile, per la quale $\delta Q = 0$, otteniamo $0 < S(B) - S(A)$, vale a dire che $S(B) > S(A)$. Concludiamo allora che in un sistema termodinamico isolato termicamente l'entropia cresce per tutte le trasformazioni irreversibili e rimane costante solo per le trasformazioni reversibili.

Le trasformazioni che avvengono spontaneamente in natura sono allora solo quelle che aumentano l'entropia complessiva dell'universo, visto come un sistema isolato. Il calore che passa spontaneamente da un corpo caldo ad uno freddo e mai viceversa

comporta un aumento di entropia. Questo implica che l'energia si degrada. Infatti qualsiasi macchina termica per compiere lavoro richiede almeno due sorgenti a temperatura T_1 e $T_2 > T_1$ e l'efficienza massima (di Carnot) di una macchina termica, $\eta_C = 1 - T_1/T_2$, è tanto migliore quanto più grande è il rapporto fra le due temperature. Inoltre in ogni macchina termica il calore sottratto alla sorgente a temperatura maggiore non può mai essere trasformato integralmente in lavoro, ma in parte viene ceduto alla sorgente a temperatura minore, e questo processo tende a diminuire la differenza di temperatura tra le due sorgenti. Quindi, per tempi lunghi, l'entropia tende a raggiungere un valore massimo, che corrisponde per un sistema isolato ad una temperatura uniforme. In questo caso, il sistema non è più in grado di compiere alcun lavoro. Considerando l'universo nella sua globalità come uno sistema isolato, lo stato a cui tende è quello di massima entropia e di temperatura uniforme, chiamato anche stato di morte termica o morte entropica dell'universo.

2. L'entropia statistica

Uno degli scopi principali della meccanica statistica è quello di spiegare le leggi della termodinamica partendo dalle leggi della dinamica molecolare. La domanda è: date le leggi del moto e le interazioni tra le molecole, quali sono le proprietà macroscopiche della materia composta da queste molecole?

Nella seconda metà del XIX secolo, Boltzmann e Clausius tentarono di derivare la seconda legge della termodinamica dalla meccanica. Questo seguendo la linea di Maxwell, che aveva già avanzato l'idea che "la seconda legge della termodinamica ha solo una certezza statistica".

Come abbiamo discusso nella sezione precedente, in un sistema isolato termicamente l'entropia non può mai diminuire. Pertanto, l'evoluzione del sistema è tale da non diventare mai più ordinato. Una dimostrazione familiare di questo principio si ha facendo cadere una goccia d'inchiostro in un recipiente contenente acqua. L'inchiostro diffonde nel bicchiere e dopo un certo tempo si forma una miscela omogenea. Non oosserviamo mai il processo inverso, vale a dire la separazione spontanea dell'inchiostro dall'acqua. Il secondo principio della termodinamica introduce quindi una freccia del tempo.

Boltzmann ha collegato la nozione di entropia al logaritmo del numero di possibili stati microscopici (o microstati) diversi compatibili con un determinato stato macroscopico (o macrostato). Per esempio, consideriamo $N/2$ molecole bianche e $N/2$ molecole nere ($N \gg 1$) all'interno di un singolo recipiente e distinguiamo lo stato microscopico di ogni molecola in base al fatto che si trovi nella metà sinistra o destra del recipiente. È chiaro che esiste un unico stato microscopico corrispondente allo stato macroscopico "tutte le molecole bianche nella metà sinistra e tutte le molecole nere nella metà destra del recipiente", mentre ci sono molti più stati microscopici che corrispondono allo stato macroscopico "le molecole bianche e nere sono equamente distribuite tra la metà sinistra e la metà destra del recipiente". Pertanto, l'entropia, o "disordine", è molto più grande in quest'ultimo stato macroscopico. Più in generale,

possiamo pensare al microstato di un gas come individuato da posizione e velocità delle singole molecole che lo compongono, mentre lo stato macroscopico è determinato da pochi parametri quali temperatura, pressione, volume e quantità di sostanza del gas. Chiamando Ω il numero di stati microscopici a cui corrisponde un dato stato macroscopico, l'entropia statistica è proporzionale al logaritmo di tale numero

$$S = k_B \ln \Omega, \tag{5}$$

dove ln indica il logaritmo naturale e $k_B \approx 1.38 \times 10^{-23}$ J/K è una costante universale, nota come costante di Boltzmann. Se nell'esempio precedente partiamo da uno stato iniziale con le $N/2$ molecole bianche a sinistra e le $N/2$ nere a destra e poi lasciamo evolvere liberamente il gas, andremo verso una situazione dove le molecole bianche e nere si equidistribuiranno tra sinistra e destra del recipiente, con $N/4$ molecole bianche a sinistra e $N/4$ a destra, e ugualmente per le molecole nere. Questo a parte fluttuazioni, trascurabili rispetto ai valori medi quando il numero di molecole N è grande (ricordiamo che una mole di gas perfetto contiene un numero di molecole pari alla costante di Avogadro $N_A \approx 6.02 \times 10^{23}\,\mathrm{mol}^{-1}$). Questo perché abbiamo un numero di microstati molto più grande nella situazione di equidistribuzione che in quella di partenza. Quindi il conseguente aumento di entropia ha una motivazione probabilistica, l'evoluzione è cioè verso lo stato macroscopico al quale corrisponde il maggior numero di stati microscopici.

Facciamo notare che la definizione di entropia di Boltzmann presuppone che, quando un sistema è in equilibrio termodinamico, tutti gli stati microscopici che soddisfano le condizioni macroscopiche del sistema siano equiprobabili. Questo non è in generale vero se abbiamo un sistema in contatto con un ambiente esterno, che ne fissa ad esempio la temperatura. Le velocità delle molecole non sono equiprobabili ma seguono una distribuzione, detta di Maxwell-Boltzmann, determinata dalla temperatura: ad esempio l'energia cinetica media delle molecole è proporzionale alla temperatura. Per i sistemi termodinamici in cui i microstati del sistema possono non avere uguali probabilità, la generalizzazione dell'entropia di Boltzmann, chiamata entropia di Gibbs [2], è data da

$$S_G = -k_B \sum_i p_i \ln p_i, \tag{6}$$

dove p_i è la probabilità del microstato i-esimo. Si noti come la (6) si riduca alla (5) quando i microstati sono tutti uguali, $p_i = 1/\Omega$ per tutti gli $i = 1, \ldots, \Omega$.

3. Entropia, energia ed informazione

3.1. Il demone di Maxwell

Possiamo dire che la discussione sulla relazione tra entropia, energia ed informazione risale al paradosso del demone di Maxwell, introdotto da James Clerk Maxwell

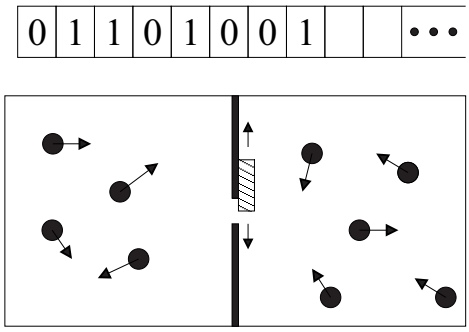


Fig. 1. – Disegno schematico del paradosso del demone di Maxwell. La memoria del demone è rappresentata da una stringa binaria (sequenza di 0 ed 1) e memorizza i risultati delle sue misurazioni delle posizioni e delle velocità delle molecole.

nel 1867 (il termine "demone" fu in realtà coniato da William Thomson, comunemente noto come lord Kelvin, nel 1874 [3]). Maxwell immaginò che un demone fosse in grado di monitorare le posizioni e le velocità delle singole molecole di un gas contenuto in una scatola e inizialmente in uno stato di equilibrio termico alla temperatura T (si veda la fig. 1). Il sistema può cioè scambiare calore ma non particelle con l'ambiente esterno, che possiamo ad esempio visualizzare come la stanza a temperatura T nella quale si trova la scatola. In questo caso il gas raggiunge uno stato di equilibrio termico alla medesima temperatura T.

All'equilibrio termico, le singole velocità molecolari sono distribuite secondo la distribuzione statistica di Maxwell-Boltzmann e con direzioni casuali. Isoliamo ora anche termicamente il gas dall'ambiente esterno. La scatola è divisa in due metà, che comunicano attraverso una porticina, che il demone apre e chiude in modo che le molecole più veloci si spostino dalla metà destra della scatola a quella sinistra e le molecole più lente vadano in senso opposto. Facendo questo molte volte, il demone separa le molecole più veloci, che finiscono sul lato sinistro della scatola, da quelle più lente, che finiscono sul lato destro. Di conseguenza, la temperatura T_s del gas a sinistra, che, secondo la teoria cinetica dei gas, è proporzionale alla velocità quadratica media delle molecole, diventa più alta della temperatura T_d del gas a destra. Poiché abbiamo ottenuto due gas a temperature diverse, possiamo ora utilizzare questa differenza di temperatura per produrre lavoro. Pertanto, il demone è stato in grado di convertire il calore fornito da un'unica sorgente a temperatura uniforme T (la stanza) in lavoro, in apparente violazione del secondo principio della termodinamica. In realtà, come vedremo più avanti, non c'è alcuna violazione, in quanto la trasformazione del calore in lavoro non è l'unico risultato del processo considerato nel suo complesso.

Il paradosso del demone di Maxwell può essere equivalentemente enunciato in termini di entropia. È chiaro che il demone introduce un ordine nel sistema: le molecole più veloci sono forzate a stare sul lato sinistro del contenitore, le molecole più lente su quello destro. Pertanto, il numero di stati microscopici accessibili al sistema diminuisce e l'entropia totale del gas si riduce, violando così apparentemente

la seconda legge della termodinamica. Inoltre, il demone potrebbe dividere la scatola in tante celle e separare con grande precisione le molecole in base alla loro velocità. Di conseguenza, l'entropia del gas diventerebbe sempre più piccola all'aumentare del numero di celle. In effetti, la nostra conoscenza dello stato microscopico del sistema aumenterebbe, il che implica una diminuzione dell'entropia, pensata come nostra ignoranza sullo stato del sistema. Tuttavia, tale violazione della seconda legge è solo apparente: infatti, come vedremo nel seguito, un'attenta analisi mostra che l'entropia totale (del gas, del demone e dell'ambiente) non diminuisce.

3.2. Il principio di Landauer

Il demone di Maxwell ha affascinato generazioni di scienziati [4], suscitando molte discussioni, e sono state proposte diverse soluzioni per risolvere il paradosso. All'inizio si riteneva che la soluzione del paradosso risiedesse nel costo energetico delle misurazioni effettuate dal demone. Per esempio, per localizzare le molecole il demone deve illuminarle e questo ha un costo energetico. Tuttavia, Rolf Landauer e Charles Bennett sono riusciti a dimostrare che il processo di misurazione può, in linea di principio, essere eseguito senza dispendio di energia [5,6]. La soluzione oggi comunemente accettata del paradosso è la seguente: i risultati delle misure devono essere immagazzinati nella memoria del demone. Semplificando la discussione, diciamo che il demone usa un bit, vale a dire il mattone fondamentale della teoria dell'informazione, che può trovarsi in due stati distinti, che chiamiamo 0 ed 1, per fissare la posizione di una molecola. Il bit sarà nello stato zero se la molecola si trova nella parte sinistra del contenitore, nello stato 1 se si trova a destra, Poichè la sua memoria è finita, il demone dovrà alla fine cancellarla per liberare le celle di memoria per le nuove misurazioni. Questa operazione comporta una dissipazione di energia Questo è il contenuto del principio di Landauer [5], enunciato nel 1961:

> *Ogni volta che un singolo bit di informazione viene cancellato, la quantità di energia dissipata nell'ambiente è almeno pari a* $k_B T \ln 2$*, dove* k_B *è la costante di Boltzmann e* T *la temperatura dell'ambiente circostante. Equivalentemente, possiamo dire che l'entropia dell'ambiente aumenta almeno di* $k_B \ln 2$*.*

Pertanto, la diminuzione dell'entropia del gas viene compensata da un aumento di entropia dell'ambiente. Infatti, per cancellare le informazioni raccolte dal demone nel processo di misurazione, dobbiamo dissipare energia nell'ambiente. Dunque, il costo energetico necessario a risolvere il paradosso, secondo il principio di Landauer, non è dovuto al processo di misurazione in sé, ma alla cancellazione delle informazioni.

Il seguente esempio è utile per illustrare il principio di Landauer. Supponiamo che un singolo bit di informazione sia rappresentato dallo stato di un sistema fisico, dato da una singola molecola in una scatola dove, come detto sopra, il bit è nello stato 0 oppure 1 se la molecola si trova a sinistra o a destra. Anche se si tratta di un sistema ad una sola molecola, possiamo applicare le leggi della termodinamica, dopo opportune medie temporali. Come è noto dal primo principio della termodinamica, $\mathrm{d}U = \delta L + \delta Q$,

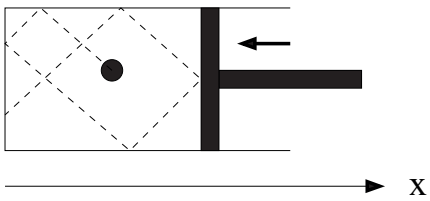


Fig. 2. – Compressione di un gas a singola molecola per mezzo di un pistone.

dove dU è la variazione dell'energia interna del gas, δL il lavoro compiuto sul gas e δQ il calore assorbito dal gas. Notiamo che usiamo d per dU e δ per δL e δQ in quanto l'energia interna (come l'entropia) è una funzione di stato, vale a dire che la variazione di energia interna dovuta ad una trasformazione termodinamica dipende solo dallo stato iniziale A e da quello finale B del sistema, mentre il lavoro e il calore non sono funzioni di stato, dipendono cioè dalla trasformazione compiuta per andare da A a B. Matematicamente, diciamo che dU è un differenziale esatto, mentre δL e δQ sono differenziali non esatti. Se consideriamo una trasformazione quasi-statica, così lenta da poter considerare il sistema come se fosse sempre in uno stato di equilibrio, possiamo scrivere, come ci si aspetta per una trasformazione reversibile, $\mathrm{d}S = \delta Q/T$, dove dS è la variazione dell'entropia del gas. Supponiamo che la nostra scatola sia a contatto termico con l'ambiente circostante, alla temperatura T, e che comprimiamo il gas per mezzo di un pistone privo di attrito (si veda la fig. 2). Se lo spostamento del pistone è dx, il lavoro compiuto sul gas è dato da

$$\delta L = -F\mathrm{d}x = -pA\mathrm{d}x = -p\mathrm{d}V, \tag{7}$$

dove F è la forza del gas sul pistone, p la sua pressione, A la superficie del pistone e V il volume del gas. Naturalmente, poiché abbiamo una sola molecola, concetti come pressione e forza devono essere intesi in un senso mediato nel tempo, vale a dire che dobbiamo calcolare una media su molte collisioni della molecola contro il pistone. Consideriamo una trasformazione che dimezzi il volume del gas. Tenendo conto dell'equazione di stato dei gas ideali, $pV = Nk_BT$, dove N è il numero di particelle presenti nel gas (qui $N = 1$), possiamo calcolare il lavoro compiuto sul gas come segue:

$$L = -\int_V^{V/2} p\,\mathrm{d}V' = -\int_V^{V/2} \frac{k_BT}{V'}\,\mathrm{d}V' = k_BT\ln 2. \tag{8}$$

Si noti che in questo caso abbiamo utilizzato il fatto che la trasformazione sia isotermica, poiché il sistema è a contatto termico con l'ambiente a temperatura T. Siccome abbiamo ipotizzato che il gas sia ideale, la sua energia interna non cambia in quanto la temperatura è costante. Pertanto, dalla prima legge della termodinamica, il lavoro compiuto sul gas si trasforma interamente in calore dissipato nell'ambiente:

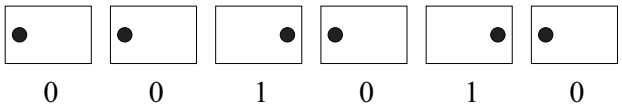


Fig. 3. – Sequenza di scatole a molecola singola associate a una stringa binaria.

$\Delta Q = -L$. Si noti che $\Delta Q < 0$ poiché il calore viene dissipato, non assorbito. La variazione dell'entropia del gas è data da

$$\Delta S = \frac{\Delta Q}{T} = \frac{-L}{T} = -k_B \ln 2. \tag{9}$$

Si ha $\Delta S < 0$ poiché, dopo la compressione, il volume disponibile per la molecola è dimezzato e quindi anche il numero di microstati disponibili si riduce di conseguenza. L'entropia del sistema diminuisce, mentre aumenta l'entropia dell'ambiente esterno al gas, ΔS_a. Poiché l'entropia totale dell'universo non può mai diminuire, si ha che $\Delta S + \Delta S_a \geq 0$. Quindi, $\Delta S_a \geq k_B \ln 2$, in accordo con il principio di Landauer.

Supponiamo ora che un messaggio binario sia memorizzato per mezzo di una sequenza di scatole a singola molecola. Ogni scatola trasporta un singolo bit di informazione, impostato nello stato 0 o 1, a seconda della posizione sinistra/destra della molecola (si veda la fig. 3). Dimostreremo ora la validità della seguente affermazione: l'informazione contenuta nel messaggio è proporzionale all'energia necessaria per cancellare il messaggio, cioè per spostare tutte le molecole sul lato sinistro (o destro) delle scatole. Rimandiamo alla prossima sezione la definizione rigorosa di contenuto di informazione di un messaggio. Ci basta per il momento notare che tale quantità può essere vista come una misura della nostra ignoranza a priori sul messaggio medesimo. È infatti chiaro che, se conoscessimo già i valori (0 o 1) dei bit, non otterremmo ulteriori informazioni dalla ricezione del messaggio. In questo caso l'informazione contenuta nel messaggio sarebbe pari a zero e non sarebbe necessario alcun dispendio di energia per cancellare il messaggio. Dimostriamo che in effetti non è necessario alcun lavoro per impostare lo stato di ogni bit a 0. Infatti, se la molecola è già nella parte sinistra della scatola, non sono necessarie altre azioni. D'altra parte, se la molecola si trova nella parte destra, possiamo spostarla a sinistra senza dispendio di energia: è sufficiente racchiuderla in una scatola più piccola, interna, e spostarla verso sinistra, come illustrato in fig. 4. Non è necessario alcun lavoro per eseguire questa operazione, poiché la molecola rimbalza altrettante volte contro la parete sinistra della scatola interna come contro quella destra. Solo nel caso in cui non conosciamo in anticipo la posizione della molecola viene trasmessa un'informazione, per cancellare la quale dobbiamo dimezzare il volume del gas e, come abbiamo visto in precedenza, questo richiede un lavoro $L = k_B T \ln 2$ da compiere sul gas, con una conseguente diminuzione di entropia $\Delta S = -k_B \ln 2$. In quest'ultimo caso, dove il contenuto informativo del messaggio è diverso da zero, dobbiamo quindi spendere energia per cancellarlo.

Per comprendere meglio la relazione tra informazione (o entropia) ed energia, è istruttivo considerare il seguente esempio, ideato da Charles Bennett, che mostra come l'informazione possa essere usata come carburante per muovere una macchina.

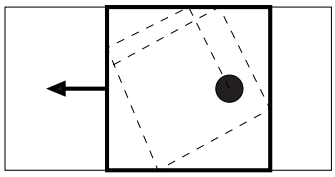


Fig. 4. – Procedura che consente il trasferimento di una molecola dal lato destro al lato sinistro di una scatola senza costi energetici.

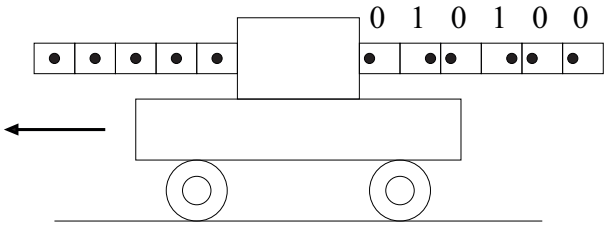


Fig. 5. – Uso dell'informazione per produrre lavoro.

Un carrello si trova in un ambiente alla temperatura T e un nastro, costituito da una sequenza di scatole contenenti una singola molecola, entra nel carrello (si veda la fig. 5). Se conoscessimo in anticipo la posizione sinistra/destra di ogni molecola, questa informazione potrebbe essere usata per estrarre il lavoro necessario a muovere il carrello. Infatti, sarebbe sufficiente inserire un pistone al centro di ogni scatola. Come mostrato in fig. 6, il pistone si sposta verso destra se la molecola si trova sul lato sinistro della scatola e verso sinistra se la molecola si trova sul lato destro. Poiché l'intero sistema è alla temperatura T, si estrae il lavoro $L = k_B T \ln 2$. Se abbiamo un nastro di N bit, il lavoro totale è $N k_B T \ln 2$ e può essere utilizzato per spostare il carrello. Sottolineiamo che, quando il nastro esce dal carrello, le molecole possono essere ovunque all'interno del volume V. Il contenuto informativo della sequenza di scatole è stato completamente perso (aumenta cioè l'entropia del sistema) e utilizzato come carburante per spostare il carrello. D'altra parte, se non conoscessimo in anticipo la posizione sinistra/destra delle singole molecole, allora non potremmo ricavare lavoro utile: infatti, se inserissimo un pistone, non sapremmo in che modo inserirlo tra le due opzioni di fig. 6. Per metà del tempo il gas produrrebbe lavoro, ma per l'altra metà del tempo il lavoro verrebbe fatto sul gas. In media, il lavoro estratto sarebbe quindi pari a zero.



Fig. 6. – Estrazione di lavoro da una singola molecola di gas, che all'inizio si trova sul lato destro o sinistro di un contenitore.

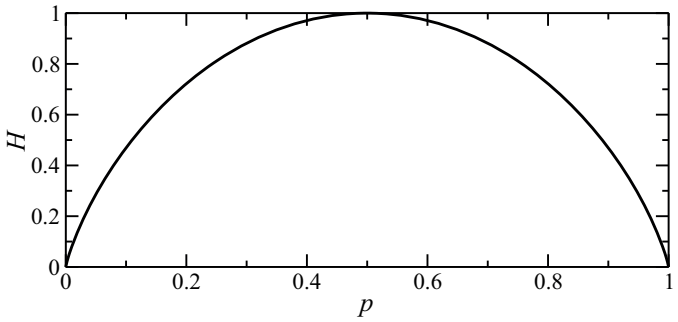


Fig. 7. – L'entropia binaria di Shannon $H(p) = -p \log_2 p - (1-p) \log_2 (1-p)$.

4. L'entropia di Shannon

Il concetto di entropia è un architrave anche della teoria dell'informazione. Il primo compito fondamentale della teoria dell'informazione è quello di quantificare l'informazione contenuta in un messaggio. Questo problema è stato risolto da Claude Shannon nel 1948 [7, 8]. Un messaggio è una sequenza di lettere scelte da un alfabeto $\mathcal{A} = \{a_1, a_2, \ldots, a_k\}$. Assumiamo che le lettere del messaggio siano statisticamente indipendenti e che la lettera a_i appaia con una probabilità a priori p_i, dove $\sum_{i=1}^{k} p_i = 1$. L'ipotesi che le lettere siano statisticamente indipendenti è stata fatta per semplificare la discussione. In pratica, questo non è il caso di molti esempi importanti. Ad esempio, in un testo il lingua italiana esistono forti correlazioni tra lettere consecutive. Tuttavia, le idee discusse in questa sezione possono essere estese a situazioni più complicate con correlazioni. Pertanto, per semplicità di seguito si assumerà sempre l'indipendenza statistica delle lettere.

L'entropia di Shannon associata alla distribuzione di probabilità $\{p_1, p_2, \ldots, p_k\}$ è definita da

$$H(p_1, p_2, \ldots, p_k) = -\sum_{i=1}^{k} p_i \log_2 p_i. \tag{10}$$

Si noti che, come usuale in teoria dell'informazione, i logaritmi sono stati scritti in base 2. L'analogia tra l'entropia di Shannon e quella statistica di Gibbs, eq. (6), è evidente: in quel caso i valori di di i individuavano il microstato di un sistema fisico. Mostriamo ora che l'entropia di Shannon quantifica quanta informazione guadagniamo, in media, quando apprendiamo il valore di una lettera del messaggio. Consideriamo il caso speciale $k = 2$ (codifica binaria) e definiamo $p_1 = p$ (dove $0 \leq p \leq 1$). Poiché $p_2 = 1 - p$, l'entropia binaria di Shannon è una funzione della sola p e possiamo scrivere

$$H_{\text{bin}}(p) \equiv H(p_1, p_2) = -p \log_2 p - (1-p) \log_2 (1-p). \tag{11}$$

Nel seguito scriveremo semplicemente $H(p)$ invece di $H_{\text{bin}}(p)$. L'entropia binaria di Shannon $H(p)$ è rappresentata nella fig. 7: è uguale a zero quando $p = 0$ o $p =$

1 e raggiunge il suo valore massimo $H = 1$ quando $p = 1/2$ (più in generale, si può dimostrare che l'entropia di Shannon $H(p_1, \ldots, p_k)$ è massima quando tutte le lettere sono equiprobabili, $p_1 = \cdots = p_k = 1/k$). Questo risultato è coerente con la nostra interpretazione di $H(p)$ come il contenuto medio di informazione di ogni lettera del messaggio. Infatti, l'informazione è una misura della nostra ignoranza a priori. Se sappiamo già che riceveremo la lettera a_1 con certezza ($p = 1$), allora non si ottiene alcuna informazione dalla ricezione di questa lettera. La stessa conclusione vale quando $p = 0$ e riceviamo sempre a_2. Se, invece, entrambe le lettere sono equiprobabili, la nostra ignoranza a priori è massima e quindi quando riceviamo una lettera otteniamo la massima informazione possibile $H(1/2) = 1$. In questo caso, si dice che abbiamo ricevuto un'unità di informazione, o bit. In genere, codifichiamo le lettere dell'alfabeto o ogni altro messaggio come sequenza di cifre binarie, cioè $a_1 = 0$ e $a_2 = 1$, e quindi il bit assume il significato di mattone fondamentale nella teoria dell'informazione.

4.1. La compressione dell'informazione

L'enorme importanza pratica della compressione dell'informazione in campi quali le telecomunicazioni o l'archiviazione dei dati è enorme. La compressione dei dati ci permette infatti di aumentare la "velocità di trasmissione" in un protocollo di comunicazione o la "capacità di memorizzazione" di un computer.

Consideriamo il seguente problema fondamentale: quanto può essere compresso un messaggio, pur ottenendo essenzialmente la stessa informazione? In altre parole, quali sono le risorse minime necessarie (in termini di uso della memoria di un computer oppure di lunghezza di una sequenza di bit inviata) per codificare un messaggio senza perderne il suo contenuto informativo? Il problema della compressione ottimale dell'informazione è strettamente legato alla ragione per la quale l'entropia di Shannon è una buona misura del contenuto di informazione di un messaggio.

Partiamo con un esempio, considerando un messaggio scritto utilizzando un alfabeto con quattro letters, $\mathcal{A} = \{a_1, a_2, a_3, a_4\}$. Assumiamo che queste lettere si presentino con probabilità $p_1 = 1/2$, $p_2 = 1/4$, $p_3 = p_4 = 1/8$. Utilizzando l'usuale codifica binaria, potremmo usare due bit per specificare le lettere dell'alfabeto:

$$a_1 \equiv 00, \quad a_2 \equiv 01, \quad a_3 \equiv 10, \quad a_4 \equiv 11. \tag{12}$$

Per inviare una lettera dell'alfabeto dobbiamo cioè inviare due bit. Visto che le lettere non sono equiprobabili, risulta però più conveniente codificarle come segue:

$$a_1 \to c_1 \equiv 0, \quad a_2 \to c_2 \equiv 10, \quad a_3 \to c_3 \equiv 110, \quad a_4 \to c_4 \equiv 111. \tag{13}$$

Per inviare una lettera codificata in questo modo abbiamo bisogno, in media, di $\sum_{i=1}^4 p_i l_i$ bit, dove l_i è la lunghezza, in bit, della lettera codificata c_i (abbiamo $l_1 = 1$, $l_2 = 2$, $l_3 = l_4 = 3$). Poiché $\sum_i p_i l_i = 7/4 < 2$, abbiamo compresso l'informazione rispetto all'usuale codifica binaria. Si noti che la buona strategia, in questo caso

come in qualsiasi altro codice di compressione utile, è quella di codificare le stringhe più probabili nelle sequenze più brevi e le stringhe meno probabili nelle sequenze più lunghe.

Shannon ha dimostrato che il tasso di compressione ottimale è dato dall'entropia di Shannon. Dato un messaggio di n lettere tratte dell'alfabeto $\mathcal{A} = \{a_1, \ldots, a_k\}$, dove ogni lettera a_i si presenta con probabilità p_i, allora, per n grande, il messaggio può essere codificato in modo affidabile usando solo $nH(p_1, \ldots, p_k)$ bit. Questo è il contenuto del teorema di codifica di Shannon:

> *Dato un messaggio in cui le n lettere sono state scelte in modo indipendente dall'alfabeto* $\mathcal{A} = \{a_1, \ldots, a_k\}$ *con probabilità a priori* $p_1, \ldots, p_k$*, esiste, per* $n \to \infty$*, un codice ottimale e affidabile che comprime il messaggio a* $H(p_1, \ldots, p_k)$ *bit per lettera.*

Si noti che, nell'esempio considerato sopra, il tasso di compressione ottimale è $H(p_1, p_2, p_3, p_4) = -\sum_{i=1}^{4} p_i \log_2 p_i = 7/4$. Dal momento che $\sum_i p_i l_i = 7/4 = H$, la compressione ottimale stabilita dal teorema di Shannon è stata raggiunta. Nessuna altra codifica potrà fare meglio della (13). Questo non è però vero in generale. Ad esempio, se le quattro lettere dell'alfabeto considerato nell'esempio sopra hanno invece probabilità $p_1 = 0.9$, $p_2 = 0.05$, $p_3 = p_4 = 0.025$, il fattore di compressione ottimale è dato da $H(p_1, p_2, p_3, p_4) \approx 0.62$, mentre la codifica (13) dà $\sum p_i l_i = 1.15$ e quindi la compressione dei dati in questo caso, sebbene utile, non è ottimale.

Se poi applichiamo la stessa codifica al caso in cui le quattro lettere sono equiprobabili, $p_i = 1/4$ per $i = 1, \ldots, 4$, otteniamo $\sum_i p_i l_i = 2,25 > 2$, mentre $H = 2$ e quindi il codice è in questo caso dannoso per l'efficienza della trasmissione dei dati. Notiamo che, siccome $H = 2$, nessuna compressione è possibile.

Il teorema di codifica di Shannon è asintotico (per messaggi con numero di lettere $n \to \infty$), ed è chiaro che comprimere sequenze molto lunghe non è molto pratico. Fortunatamente, esistono metodi piuttosto efficienti per codificare sequenze corte. Consideriamo ad esempio il codice di Huffman, mostrato nella tabella I, dove si usa un alfabeto binario $\{0, 1\}$ e la procedura di codifica viene applicata a sequenze di quattro bit. Esistono $2^4 = 16$ stringhe di questo tipo ($0 \equiv 0000, 1 \equiv 0001, \ldots, 15 \equiv 1111$). Sia P_i la probabilità che si verifichi la sequenza i, con $i = 0, \ldots, 15$. Abbiamo $P_0 = p_0^4, P_1 = p_0^3 p_1, \ldots, P_{15} = p_1^4$. Se consideriamo, per esempio, il caso con $p_0 = 3/4$ e $p_1 = 1/4$, la migliore compressione possibile per un messaggio di quattro lettere è data da $4H(p_0, p_1) \approx 3.25$, mentre il codice di Huffman dà in media $\sum_{i=0}^{15} P_i l_i \approx 3.27$ bit, un valore molto vicino a quello ottimale. Questo esempio mostra la potenza dei codici di compressione dei dati.

La possibilità di comprimere dati è una caratteristica generale per un messaggio. Il teorema di Shannon infatti ci dice che, nella misura in cui le lettere di un messaggio non sono equiprobabili, la compressione dei dati è possibile. Questo è il caso, per esempio, per le lettere di un qualsiasi testo (in italiano, in inglese o in un'altra lingua), le quali non sono equiprobabili ma appaiono con frequenze diverse.

TABELLA I. – *Codifica di Huffman, per* $p_0 = 3/4$, $p_1 = 1/4$.

Messaggio	Codifica di Huffman
0000	10
0001	000
0010	001
0011	11000
0100	010
0101	11001
0110	11010
0111	1111000
1000	011
1001	11011
1010	11100
1011	111111
1100	11101
1101	111110
1110	111101
1111	1111001

5. Osservazioni conclusive

L'entropia non è solmente un concetto chiave nella termodinamica e nella fisica statistica, ma ha assunto un ruolo fondamentale in svariati campi, tra i quali la teoria dell'informazione. Più in generale, il concetto di entropia influenza la nostra percezione dell'universo e in particolare dello sviluppo tecnologico. Ogni processo produttivo provoca un aumento dell'entropia, riducendo allo stesso tempo l'energia utile disponibile. Lo sviluppo sostenibile può rallentare il processo di crescita entropica, ma non annullarlo completamente, e questo limite all'uso delle risorse è una realtà con cui l'uomo deve imparare a convivere.

Notiamo infine che non si sono qui discussi gli aspetti dell'entropia in relazione ai sistemi dinamici e alla meccanica quantistica. Nel primo caso, l'entropia dinamica (di Kolmogorov-Sinai) quantifica in modo rigoroso la complessità di un sistema dinamico [9, 10]. Questo è un problema di grande rilevanza concettuale ma anche pratica, legato al problema di conoscere le minime risorse computazionali richieste nella simulazione di modelli dinamici che descrivano l'evoluzione di sistemi complessi, di interesse non solo per la fisica e la matematica, ma anche per una classe molto più ampia di problemi in campi diversi quali la chimica, la biologia, l'economia, la medicina, l'ingegneria, le scienze sociali, la meteorologia, la dinamica delle popolazioni e così via. L'estensione del concetto di entropia al regime quantistico (tramite l'entropia di von Neumann) è di particolare importanza in relazione alla termodinamica dei sistemi alla nanoscala, alla teoria dell'informazione quantistica e alla misura delle correlazioni puramente quantistiche, note come entanglement, di interesse come risorsa per lo sviluppo delle tecnologie quantistice [11].

Bibliografia

[1] Clausius R., *The Mechanical Theory of Heat – with its Applications to the Steam Engine and to the Physical Properties of Bodies* (J. Van Voorst, London) 1867, p. 365.
[2] Jaynes E. T., "Gibbs *vs.* Boltzmann entropies", *Am. J. Phys.*, **33** (1965) 391.
[3] Thomson W., "Kinetic Theory of the Dissipation of Energy", *Nature*, **9** (1874) 441.
[4] Rex A., "Maxwell's Demon —A Historical Review", *Entropy*, **19** (2017) 240.
[5] Landauer R., "Irreversibility and heat generation in the computing process", *IBM J. Res. Dev.*, **5** (1961) 183.
[6] Bennett C. H., "The thermodynamics of computation – A review", *Int. J. Theor. Phys.*, **21** (1982) 905; "Demons, engines and the second law", *Sci. Am.*, **257** (1987) 108.
[7] Shannon C., "A mathematical theory of communication", *Bell Syst. Tech. J.*, **27** (1948) 379; 623.
[8] Cover T. M. and Thomas J. A., *Elements of Information Theory* (John Wiley & Sons, New York) 1991.
[9] Ott R., *Chaos in Dynamical Systems*, 2nd edition (Cambridge University Press, Cambridge) 2002.
[10] Ford J., "How random is a coin toss?", *Phys. Today*, **36** (1983) issue No. 40.
[11] Benenti G., Casati G., Rossini D. and Strini G., *Principles of Quantum Computation and Information (A Comprehensive Textbook)* (World Scientific, Singapore) 2019.