Allegato 1.

Alcuni concetti di Fisica delle Reti Complesse.

a1.1 Universalità.

In *meccanica statistica*, l'universalità è l'osservazione che ci sono propietà, per una ampia classe di sistemi, che sono indipendenti dai dettagli dinamici del sistema.
I sistemi che mostrano universalità tendono ad essere *caotici* e spesso hanno un gran numero di parti interagenti tra loro.
La nozione di universalità è nata dallo studio delle *transizioni di fase* attraverso la meccanica statistica. Le transizioni di fase sono caratterizzate da un *parametro d'ordine*, come la densità o la magnetizzazione, che cambiano in funzione di un *parametro del sistema*, ad esempio la temperatura.

Lo speciale valore del parametro al quale il sistema cambia la sua fase è il *punto critico del sistema*; per sistemi che esibiscono universalità il valore più vicino a tale parametro è il *valore critico*, mentre, per quelli meno sensibili il parametro d'ordine dipende dai dettagli del sistema.
Se il parametro β è critico al valore β_c, il parametro d'ordine è approssimato dalla relazione:

$$a = a_0 ||\beta - \beta_c||^{\alpha} \text{ (a.1)}$$

L'esponente α è un esponente critico del sistema. *La rilevante scoperta fatta nella seconda metà del ventesimo secolo è che sistemi molto differenti tra loro hanno gli stessi esponenti critici, quindi universalità.*
Esistono però differenti *classi di universalità*, ossia classi caratterizzate da diversi esponenti critici a cui i sistemi fisici in esse ricadenti obbediscono.
L'universalità è anche osservata in sistemi non in equilibrio, come sitemi di particelle interagenti, modelli di reazione-diffusione, o sistemi auto organizzati.

Un esempio di universalità attinente alle reti fluviali, ma in generale estendibile a moltissime altre reti di trasporto, è quella del sistema circolatorio umano; l'analogia con i corsi d'acqua è stata introdotta da *Rinaldo* coniando il curioso termine "*hydrologic allometry*" [*Rinaldo et al.*, 2006]. Il concetto di efficienza di una rete è già stato introdotto nella descrizione delle OCNs, pertanto qui si è riporta direttamente il risultato che interessa per spiegare l'universalità manifestata attraverso la legge a potenza, detta *relazione di Kleiber*, del sistema di trasporto e consegna dei nutrienti ad un essere vivente (animale e essere umano), dove Mè la massa corporea, e Bè il tasso metabolico basale.

La legge è:

$$B \propto M^{3/4}, \text{ (a.2)}$$

la quale discende dalle relazioni tra il volume totale di sangue Cutilizzato (pari alla somma del flusso di nutrienti) ed il tasso metabolico totale B, con la lunghezza caratteristica Ldell'organismo, nello spazio Euclideo, quindi per $D = 3$ [*Banavar et al.*, 1999, 2002].

$$\left.\begin{array}{l} B \propto L^D \\ C = M \propto L^{D+1} \text{per grafo ottimizzato} \end{array}\right\} \rightarrow B = M^{\frac{D}{D+1}}$$

(a.3)

Questa relazione è sorprendentemente valida per M che spazia in un intervallo di oltre 21 ordini di grandezza, dagli organismi unicellulari ai più grandi mammiferi del pianeta. La spiegazione di tale universalità è attualmente molto dibattuta [*West et al.*, 1999; *Banavar et al.*, 1999, 2002; *Painter et al.*, 2000; *Dodds et al.*, 2003], quello che però sembra assodato è il concetto che il trasporto dei nutrienti necessari alla vita debba avvenire con la quantità minima possibile trasportata di materia, nella fattispecie del sistema circolatorio il sangue.

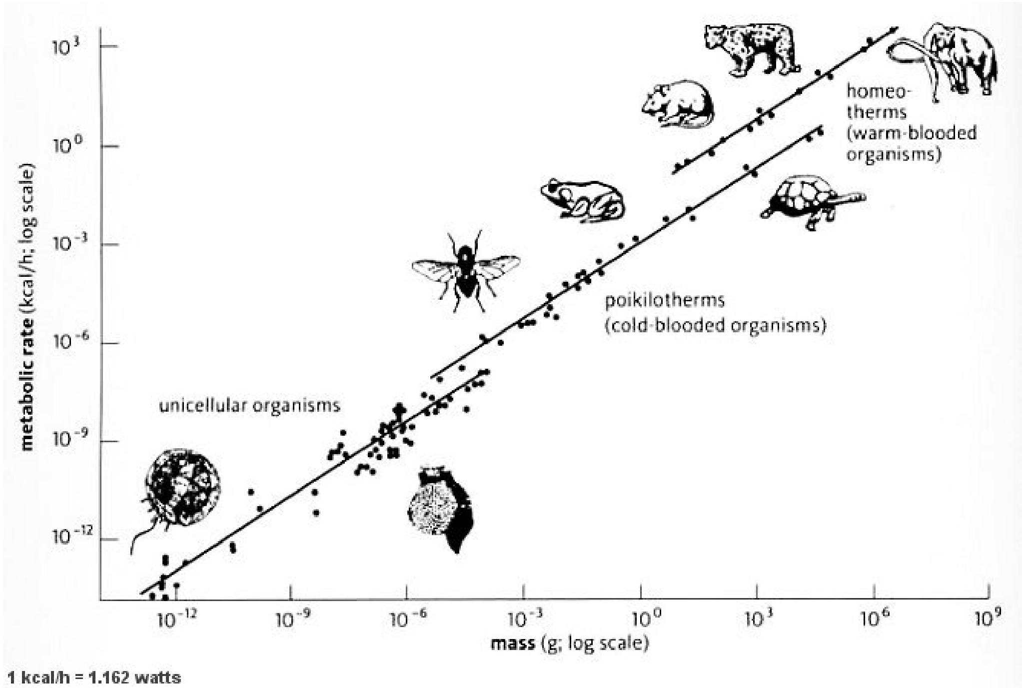


Figura a1.1 [da `http://universe-review.ca/R10-35-metabolic.html`]: rappresentazione bi-logaritmica della legge di *Kleiber*, $lnB = \frac{3}{4} lnM$; in precedenza ad essa si pensava che l'esponente fosse pari a 2/3.

L'universalità, tuttavia, è anche presente nelle statistiche di altre quantità anche se esse non sono distribuite a legge di potenza.
Per esempio in un'ampia varietà di casi la distribuzione $P(x)$ della variabile x non è invariante alla scala di analisi (*scale invariant*), piuttosto è concentrata attorno a piccoli valori come 4, 5, 6 tipicamente. Questo è chiamato *Small World Effect* ed è stato studiato soprattutto per le reti sociali dove nei grafi i vertici rappresentano gli individui e i lati-collegamenti un piccolo numero di relazioni tra gli individui stessi.
L'effetto *Small World* è stato inserito nello *Small World Model* [*Watts et al.*, 1998, 1999] il quale descrive particolari grafi il cui diametro rimane molto piccolo mentre il numero dei vertici aumenta. Tale modello è risultato essere ottimamente aderente al WWW, alle reti sociali e alle reti di coautori dove i diversi scienziati scrivono la pubblicazione congiuntamente [*Newman*, 2001a, 2001b].

Un'estensiva ricerca in ambito dello *Small World Model* applicato alle reti sociali è attualmente in corso alla Columbia University nella quale si vuole anche verificare la validità della cosidddetta teoria "dei sei gradi di separazione" [*Watts*, 2003].

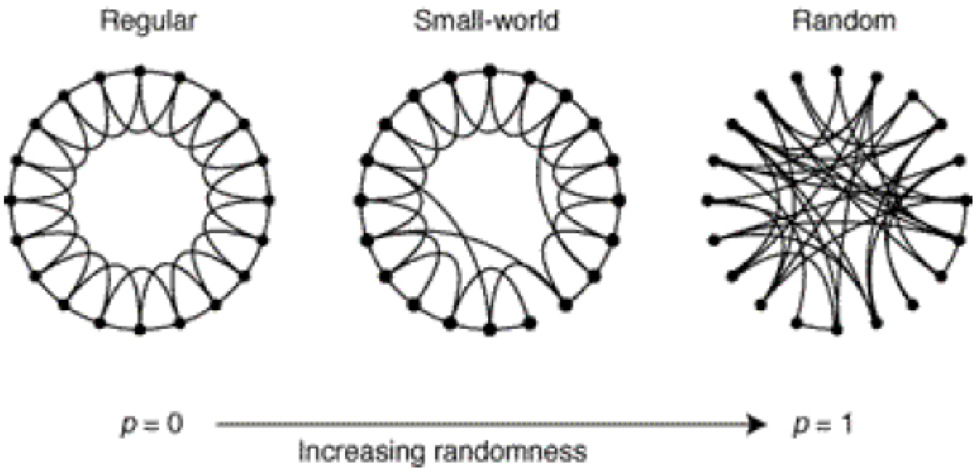


Figura a1.2 [*Colizza*, 2004]: il modello *Small World* interpola bene i modelli di rete estremi, il primo completamente ordinato e non "*clusterizzato*", l'ultimo completamente casuale e "*clusterizzato*"; il numero di nodi $n = 20$ e il numero di collegamenti, di lunghezza l, è costante, con $p = \langle k \rangle / n = \langle C_{rand} \rangle$ (≡coefficiente di *clustering* medio≡probabilità di connessione) e $\langle k \rangle = 2\langle l \rangle / n = np$per n grande (≡grado medio di *Erdös-Rényi*)

a1.2 Leggi di potenza e distribuzioni a scala libera.

Quando la probabilità di misurare un particolare valore di qualche quantità varia inversamente come una potenza di quel valore, la quantità è detta seguire una legge a potenza (*power law*), anche conosciuta come *legge di Zipf's* o *distribuzione di Pareto* (in realtà le due leggi differiscono nella rappresentazione grafica della probabilità cumulata $P(x)$, il grafico di Zipf pone $P(x)$ in ordinata e x in ascissa, quello di Pareto l'opposto).

La forma della distibuzione di probabilità che segue una legge a potenza è generalmente descritta dalla seguente relazione:

$$p(x) = Cx^{-\alpha}, \text{ (a.4)}$$

da cui si ricava la relazione logaritmica,

$$log[p(x)] = -\alpha log(x) + log(C). \text{ (a.5)}$$

La costante C nell'equazione (a.4) è fornita dalla richiesta di normalizzazione, ossia

$$\int_{x_{min}}^{\infty} p(x)dx = 1 \rightarrow C = (\alpha - 1)x_{min}^{\alpha-1}, \text{ (a.6)}$$

nella quale è immediatamente visibile che deve essere $\alpha > 1$ altrimenti l'equazione diverge; le leggi a potenza con esponente minore dell'unità non possono essere normalizzate e normalmente non si verificano in natura.

La frequenza degli eventi o gli effetti di variazione di dimensione in sitemi critici auto organizzati (SOCS), per esempio la legge di Gutenberg-Richter dell'intensità dei terremoti, e le leggi di Horton descriventi i sitemi fluviali, segueno una legge a potenza.
Tuttavia leggi a potenza sono osservate in svariati campi, includendo la fisica, biologia, geografia, economia, linguistica, le guerre ed il terrorismo.

Le leggi a potenza sono tra le più frequenti leggi di scala che descrivono l'invarianza di scala trovabile in molti fenomeni naturali.

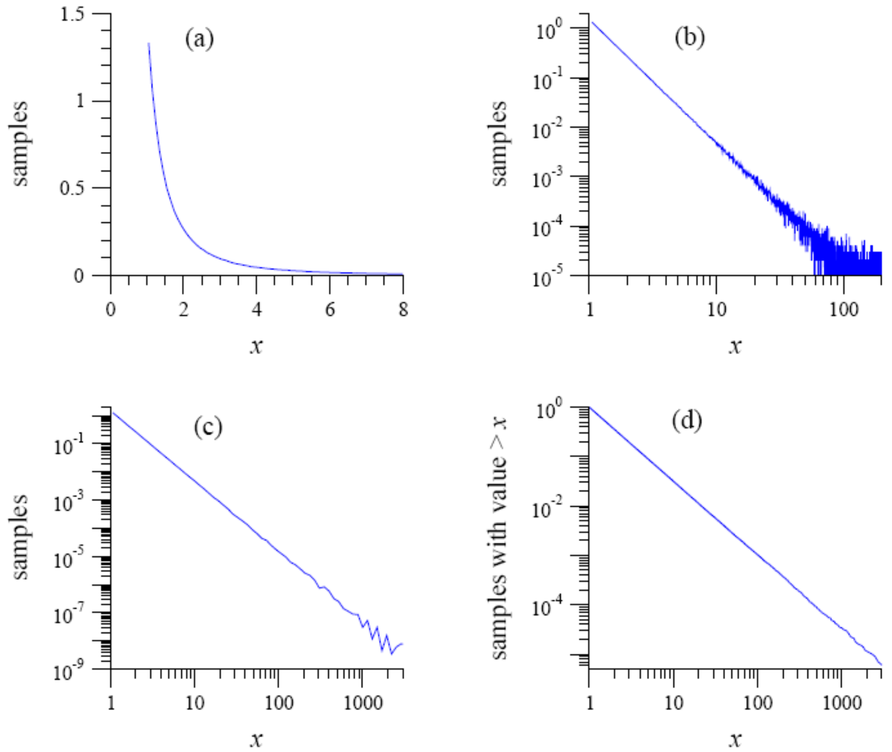


Figura a1.3 [*Newman*, 2005]: (a) istogramma di un insieme di un milione di numeri casuali, il quale ha una distribuzione a legge di potenza con esponente $\alpha = 2.5$; (b) lo stesso istogramma in scala logaritmica dove si notano le fluttuzioni statistiche nella coda; (c) istogramma costruito con la tecnica del "*logarithmic binning*" ossia normalizzando il numero di campioni nell'intervallo in cui ricadono; (d) istogramma della probabilità cumulata $P(x)$la quale segue anch'essa una legge a potenza ma con esponente $\alpha - 1 = 1.5$.

Una distribuzione a legge di potenza è anche alle volte chiamata *distribuzione* "*a scala libera*". Questo perchè la legge di potenza è la sola distribuzione che è la stessa a qualsiasi scala la si guardi[1].

Sistemi a scala libera sono più in generale quei sistemi nei quali la singola

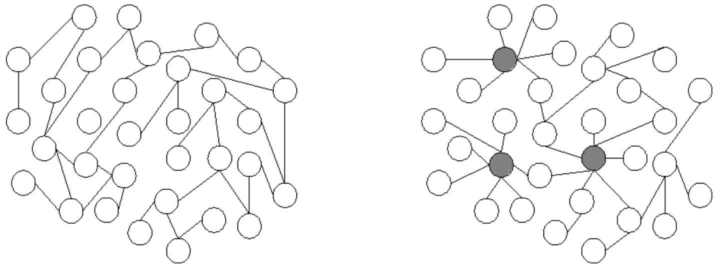
[1] Poichè sia la legge di potenza sia la distribuzione log-normale sono distribuzioni asintotiche, può essere facile confonderle non utilizzando metodi statistici robusti come il modello Bayesiano di selezione o test sulle ipotesi statistiche. Un grafico bi-logaritmico di una distribuzione log-normale può spesso sembrare dritto per certi rapporti di x e y. Una regola per verificare se la distribuzione è conforme a una *power law* è quella di verificare se, in un grafico bi-logaritmico, risulta rettilinea a tre o più ordini di grandezza ("ingrandimento").

lunghezza macroscopica (definita spesso come *lunghezza di correlazione*), la dimensione di scala, o la scala di tempo che li caratterizza, diverge lasciandoli senza scala.

Tale situazione è tipica dei fenomeni critici e produce una legge a potenza seguita dalle grandezze fisiche macroscopiche del sistema.
Il preciso punto al quale la lunghezza di scala diverge è chiamato *punto critico* o *punto di transizione di fase continua*; si è quindi associato il concetto di universalità al concetto di legge a potenza, ossia l'universalità si esibisce attraverso una legge a potenza ed ogni classe di universalità ha un "ben preciso" esponente critico. Abbiamo visto però come esistono delle eccezioni a tale regola, vedi ad esempio il modello delle *Small World Networks* descritto sinteticamente nella sezione precedente.

Attenzione però che non tutto ciò che esibisce una legge a potenza ha a che fare con complessità o criticità. Esistono meccanismi come i processi moltiplicativi, le combinazioni di esponenziali, le inversioni di quantità, e i cammini aleatori, i quali sono esempi di alcuni metodi di generazione di leggi a potenza [*Caldarelli*, 2005; *Newman*, 2005].

Recentemente [*Barabasi et al*, 1999] il concetto di distribuzione a scala libera è stato associato al concetto di reti a scala libera (*scale free networks*) introdotte da Barabasi studiando dapprima la topologia del World Wide Web e poi applicando li stessi concetti ad alcune reti sociali e biologiche [vedi anche *Vázquez et al.*, 2003].
Una *scale free network* è una struttura con distribuzione a scala libera ma non più caratterizzata da una connettività puramente aleatoria; sono presenti nodi fulcro (*hubs*) i quali hanno più collegamenti in media di tutti gli altri nodi della rete, formati da un principio di "attacco preferenziale" [*Barabasi et al.*, 2000].



(a) Random network (b) Scale-free network

Figura a1.4 [*Castillo*, 2004]: (a) rete random; (b) rete a scala libera nella quale i nodi-fulcro sono evidenziati; ogni grafo ha 32 nodi e 32 collegamenti ma entrambi sono stati realizzati per apparire in modo chiaro nel piano, quindi nemmeno la (a) è una struttura puramente casuale.

a1.3 Effetti di scala finita (*Finite Size Effects*).

Ricordando la natura discreta della funzione distribuzione di probabilità nella pratica, è da sottolineare che anche per una funzione a scala libera c'è una tipica scala che è sempre presente, dovuta alla dimensione finita dei campioni analizzati.

Nel caso specifico delle reti si ha un grado minimo che è uno (non si può avere meno di un collegamento in un *grafo*) e un grado massimo che è la dimensione della rete (non si possono avere più di $n-1$ collegamenti). Come nel caso del World Wide Web il grafo è molto ampio ed è possibile spostare il *cutoff superiore* lungo l'asse x delle ascisse quanto si vuole.
Si ha però una deviazione dal comportamento a legge di potenza quando x diventa troppo simile alla dimensione del sistema e quindi si hanno errori nella determinazione degli esponenti.

Risulta quindi controindicato prendere campioni di dati (*datasets*) troppo grandi poichè in tal caso la media, o momento primo, calcolata come sotto indicato (Formula 10), diverge senza limiti. La distribuzione in tal caso si dice non avere media finita, ed è tipica per valori di $\alpha \leq 2$ come accade per la distribuzione delle guerre o dei pennacchi solari.

$$\langle x \rangle = \int_{x_{\min}}^{\infty} x p(x) \mathrm{d}x = C \int_{x_{\min}}^{\infty} x^{-\alpha+1} \mathrm{d}x = \frac{C}{2-\alpha} [x^{-\alpha+2}]_{x_{\min}}^{\infty} \quad \text{(a.7)}$$

La relazione () è utilizzabile per un dataset finito e quindi la si utilizza quando la coda della distribuzione viene tagliata (*tail cutoff*). In tal caso per ricavare l'esponente α non si può più procedere con la forma analitica tradizionale di seguito riportata (Formula 11), ma con altre tecniche abbastanza elaborate basate sulla funzione *beta* di Legendre [*Newman*, 2005].

$$\alpha = 1 + n\left[\sum_{i=1}^{n} \ln \frac{x_i}{x_{min}}\right] - 1 \text{ (a.8)}$$

Spesso si pone anche il problema di stabilire un valore x_{min}, infatti poche distribuzioni reali seguono una legge a potenza in tutto il campo delle ascisse, ed in particolare per valori piccoli della variabile misurata.
Infatti la legge della distribuzione a potenza $p(x) = Cx^{-\alpha}$ diverge per $x \to 0$ con $\alpha > 0$. In tali casi l'esponente della legge viene calcolato omettendo la regione di valori compresi tra l'origine e x_{min} e la distribuzione viene detta "a coda" (*power law tail distribution*).

In conclusione la traduzione più adatta di *finite-size scaling* sembra essere quella di *leggi di scala per sistemi "a taglia fine"*, ossia leggi di potenza per campioni non troppo numerosi.

E' evidente come trovare esattamente li esponenti sia complesso, cosi come lo sono altre rilevanti questioni relative alle reti complesse, si rimanda pertanto a pubblicazioni specifiche [*Caldarelli*, 2005; *Newman*, 2005; *Albert et al.*, 2002].