

fare misure abbastanza precise sullo stato del tempo oggi (le condizioni iniziali del sistema), da queste si può ricavare una previsione a breve termine, ma con il passare del tempo le previsioni diventano sempre meno affidabili. Analogamente prevedere dove si troverà una certa boccia del biliardo dopo 1 rimbalzo è relativamente semplice ma se ci proponiamo di prevedere dove si trovi dopo 20 rimbalzi sulle altre bocce la cosa è praticamente impossibile. Una persona incaricata di descrivere il comportamento del tempo in una certa città per lunghi periodi, non potrà fare altro che fare una lista delle condizioni meteorologiche che incontra ogni giorno, senza la possibilità di trovare una regola semplice per spiegarle in blocco (tipo se oggi piove domani c'è il sole).

In matematica ci sono teoremi che assicurano che in molti casi  e  appaiono necessariamente insieme: l'uno implica l'altro.

La ragione "pratica" di ciò è che a causa dell'amplificarsi degli errori iniziali conoscere il comportamento di un sistema per lungo tempo "equivale" a conoscere la condizione iniziale con una grande precisione, grande precisione significa conoscere i dati con molte cifre decimali e quindi molti bits di informazione sono necessari.

Inoltre più aumenta il tempo per cui è richiesta la conoscenza del comportamento del sistema, più la conoscenza della precisione iniziale deve essere accurata.

Ovviamente tradurre questo semplice ragionamento euristico equalitativo in enunciati precisi e quantitativi e dimostrazioni matematiche non è un compito facile e richiede l'uso di strumenti raffinati della matematica contemporanea che noi in questa esposizione tratteremo in maniera informale ed in casi particolari.

Nel seguito vedremo una verifica "sperimentale" di questo principio in una famiglia di casi interessanti.

Le seguenti figure rappresentano la complessità e la sensibilità rispetto al dato iniziale di una famiglia di sistemi. Nelle figure ogni punto rappresenta un sistema, nella figura 2 il punto viene colorato a seconda che il sistema corrispondente sia più o meno sensibile al dato iniziale (gli errori si amplificano velocemente), mentre nella figura 1 il punto viene colorato a seconda che il comportamento del sistema sia più o meno complesso.

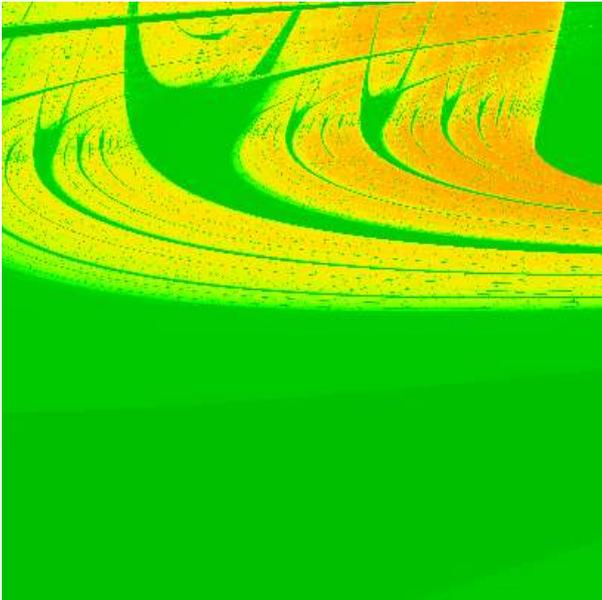


Figura 1(Complessita)

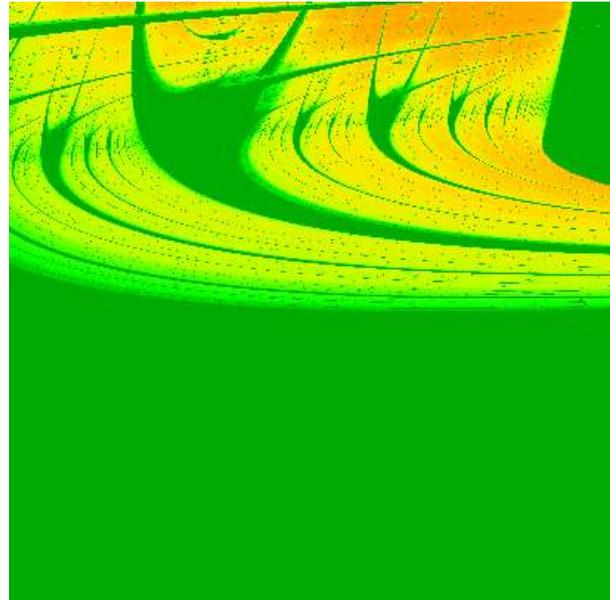


Figura 2 (Instabilita)

Come si vede le due figure sono molto simili. Questa in un certo senso e'una verifica sperimentale della stretta relazione che c'e' fra  e .

Ma cosa rappresentano esattamente le figure?

Per rispondere a questo introdurremo un modo con cui si puo misurare quantitativamente la complessita e l'instabilita, quindi abbiamo bisogno di precisare qualisono gli oggetti con cui andremo a lavorare.

Cosa e' un sistema dinamico in matematica

Fino ad adesso abbiamo parlato di sistemi o di dinamica senza specificare di preciso che cosa si intende.

In matematica, la prima cosa che si deve fare e' quella di dare definizioni precise degli oggetti di cui si parla.

Se si costruiscono dei modelli allo scopo di studiare un certo fenomeno spesso conviene considerare il modello piu semplice che presenta il detto fenomeno, in modo da eliminare complicazioni inutili. Per studiare i misteri della dinamica basta un modello semplicissimo che e' fatto di due oggetti:uno spazio e una funzione.

Quando si parla di dinamica di un sistema (o di dinamica) si pensa al fatto che il sistema parte con una condizione iniziale (in genere conosciuta in modo piu o meno preciso) e poi si evolve nel tempo.

Il caso piu semplice di questo tipo di sistemi e' quello che in matematica si chiama sistema dinamico a tempo discreto. E' una cosa semplice che da luogo a comportamenti estremamente complicati e affascinanti.

La formalizzazione del concetto di sistema dinamico e' la seguente: un sistema dinamico e' una funzione che agisce su un certo spazio, ovvero si ha uno spazio X e una funzione f da X in X . Ricordiamo che una funzione f e' un modo di associare ad ogni punto x di X un altro punto $y=f(x)$ di

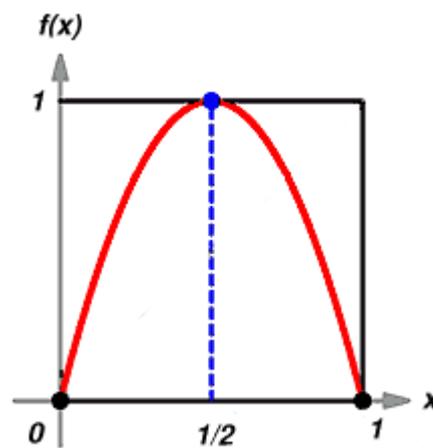
X .

Data una condizione iniziale x la dinamica si ottiene *iterando* la funzione f . Cioe si parte da x , poi si va in $f(x)$ (che e' un altro punto del nostro spazio al quale si puo applicare la f), quindi si va in $f(f(x))$ e cosi via generando una traiettoria lunga quanto si vuole.

Ad esempio

Facciamo un esempio: consideriamo come spazio X l'insieme dei numeri fra 0 e 1, estremi compresi (in notazione matematica si scrive $X=[0,1]$) quindi i punti x del nostro spazio saranno i numeri compresi fra 0 e 1. Consideriamo la funzione f definita da $f(x)=4x(1-x)$. Questa funzione associa un numero in $[0,1]$ ad un numero in $[0,1]$.

Ad esempio al numero $x_0=1/2$ viene associato il numero $x_1=f(1/2)=4*1/2*1/2=1$. Nella seguente figura e' illustrato il grafico di f .



La dinamica del sistema si ottiene *iterando* la f , cioe se la condizione iniziale (al tempo 0) e' un certo x_0 , allora al tempo 1 avremo $x_1=f(x_0)$, e poi $x_2=f(x_1)$, e cosi via $x_n=f(x_{n-1})$ per ogni n . Equivalentemente si avra' $x_2=f(f(x_0))$, $x_3=f(f(f(x_0)))$, $x_n=f(f(\dots n \text{ volte} \dots(x_0)\dots))$.

Ad esempio la traiettoria che parte dal punto $x_0=1/2$ sara' la seguente:
 $x_1=f(1/2)=1$, $x_2=f(1)=0$, $x_3=f(0)=0$, $x_4=f(0)\dots$

e cosi via ottenendo sempre 0 (cioe $x_n=0$ per tutti gli $n \geq 3$) perche si continuera' ad avere $f(0)=0$. In generale si dicono punti fissi i punti x per cui $f(x)=x$. Quando la dinamica arriva in uno di questi vi si ferma per sempre.

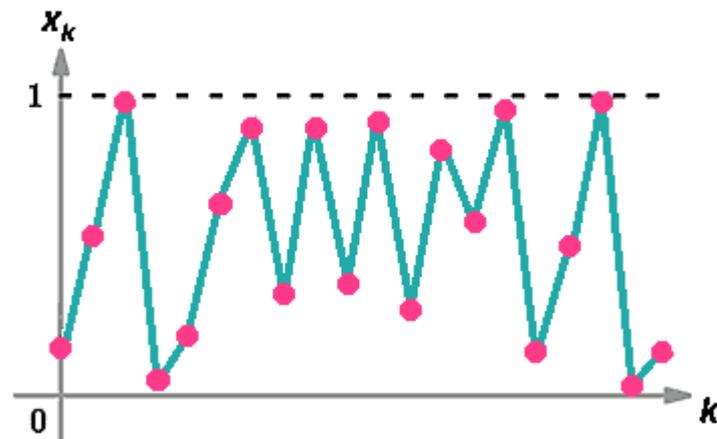
Se facciamo partire la dinamica da altri punti, diversi da $1/2$ in generale potranno accadere cose diverse, per esempio la traiettoria potrebbe convergere verso un'orbita periodica, cioe gli x_n

ripeteranno periodicamente gli stessi valori (ad es ... $x_{10}=0.3, x_{11}=0.6, x_{12}=0.3, x_{13}=0.6, \dots$).

In altri casi si potra avere una successione di valori che sembrano uscire a "casaccio" come ad esempio se $x_0=0.1$ si avra' $x_1=0.36$, $x_2=0.64$, $x_3=0.9216$, $x_4=0.289013\dots$, $x_5=0.82193\dots$.

Se disegniamo un grafico di questa traiettoria otterremo una figura simile alla seguente. Anche se l'apparenza e' quella di una successione di punti a caso quella che e' disegnata qui sotto e' un'orbita

deterministica: ogni punto e' determinato dal precedente (mediante l'applicazione della funzione f).

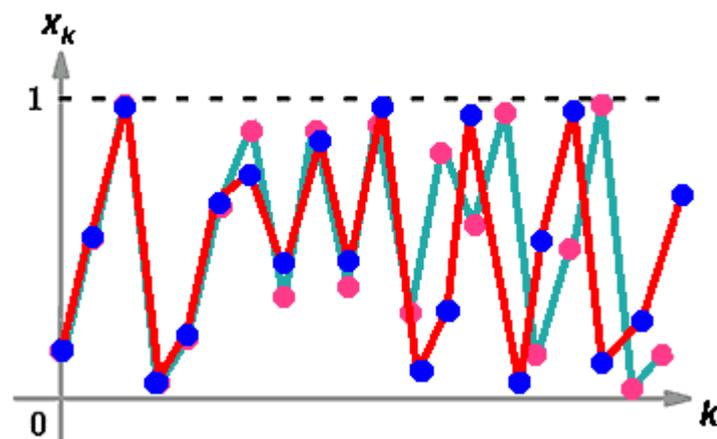


L'errore iniziale

Vediamo piu approfonditamente il punto  nel nostro esempio.

Vediamo ad esempio come un piccolo errore iniziale viene amplificato nei passi successivi della dinamica.

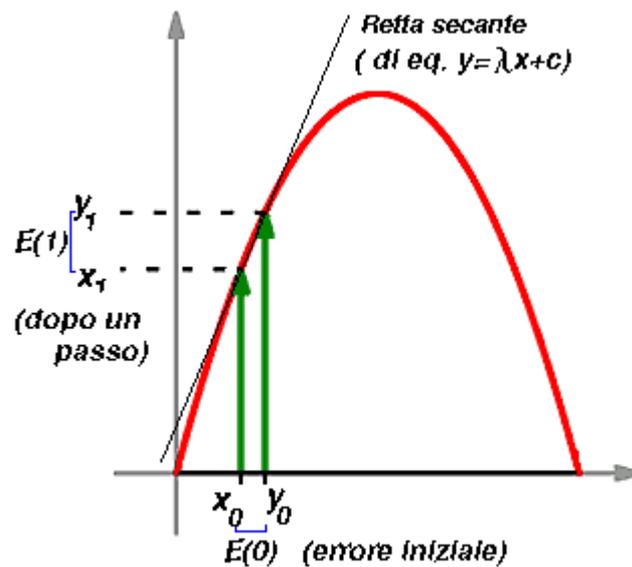
Nella figura sotto e' riportata anche la dinamica (indicata con y_n) di un punto che parte con una condizione iniziale y_0 leggermente maggiore di x_0 . Come si vede la distanza fra i passi successivi x_n, y_n della dinamica aumenta fino ad essere confrontabile con la dimensione dello spazio stesso, per poi fluttuare casualmente (e' ovvio che i due punti non possono allontanarsi indefinitamente, visto che l'intero spazio che consideriamo ha diametro 1).



Vediamo piu in dettaglio come si genera questo allontanarsi di orbite che partono inizialmente vicine.

Nella figura si vede che due condizioni iniziali che partono vicine, con un errore $E(0)$ vengono

mandate dalla f in condizioni finali che distano $E(1)$. Dalla figura si vede che $E(1) = \lambda E(0)$, ovvero l'errore finale e' dato dall'errore iniziale moltiplicato per il coefficiente angolare della secante al grafico passante per x_0 e y_0 .



Nella figura il coefficiente angolare e' piu o meno 2, dunque in questo caso dopo un passo l'errore si raddoppia.

Chi dei lettori ha seguito un corso di analisi matematica ricordera' che il coefficiente angolare della retta secante di due punti molto vicini ha a che fare con un concetto molto famoso: [la derivata](#).

Per la precisione esiste un teorema che ci assicura che $E(1)/E(0) = f'(\xi)$ dove ξ e' un punto compreso fra x_0 e y_0 . Se x_0 e y_0 sono molto vicini anche ξ e' molto vicino a x_0 e siccome la derivata e' una funzione continua si ha che il suo valore in x_0 e' molto vicino al valore calcolato in ξ e quindi possiamo dire che approssimativamente si ha che l'errore iniziale viene moltiplicato per (il valore assoluto della) derivata della funzione in x_0 cioe

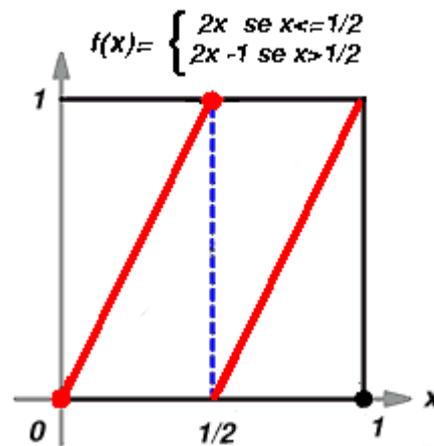
$$E(1)/E(0) = |f'(x_0)|.$$

Ad esempio**Se la derivata e' sempre grande?**

Il lettore potra' facilmente intuire che se consideriamo un sistema con $f'(x) > a > 1$ per ogni x in $[0,1]$ allora punti che partono inizialmente molto vicini si allontanano a velocita' esponenziale perche' ogni volta l'errore sara' moltiplicato per un numero maggiore di a ,

dunque $E(n) \geq E(0) \cdot a \cdot a \cdot a \dots n \text{ volte} \geq E(0) \cdot a^n$.

Questo succede ad esempio nel sistema rappresentato nella seguente figura



Qui infatti $f'(x) = 2$ per "quasi" ogni x . Quindi piccoli errori iniziali vengono raddoppiati ad ogni passo.

Il coefficiente di Lyapunov: una misura quantitativa della velocita di amplificazione dell'errore iniziale.

In generale si avra' che se x_1, x_2, \dots, x_n e' la traiettoria che parte da x_0 l'errore al passo n sara' approssimativamente $E(n) = E(0) \cdot |f'(x_0) \cdot f'(x_1) \cdot \dots \cdot f'(x_n)|$ moltiplicando l'errore iniziale per i valori assoluti delle derivate nei punti incontrati durante la dinamica. Siccome gli

$f'(x_0), f'(x_1), \dots, f'(x_n)$ potranno essere tutti diversi (anche minori di 1) l'errore iniziale viene moltiplicato per una successione di numeri diversi.

Comestimare allora la velocita' di allontanamento di traiettorie che partono vicine?

Ci possiamo chiedere "in media" per quanto e' moltiplicato l'errore ad ogni passo.

Per rispondere a questa domanda immaginiamo che esista un tale coefficiente di dilatazione medio.

Allora si potrà scrivere $E(n) = E(0) * (f^{medico})^n = E(0) * |f'(x_0) * \dots * f'(x_n)|$ da cui

$$E(n) = E(0) * (f^{medico})^n = E(0) * |f'(x_0) * \dots * f'(x_n)| \text{ e}$$

$$\log((f^{medico})^n) = \log|f'(x_0) * \dots * f'(x_n)| + \text{costante}$$

$$\log(f^{medico}) = [\log|f'(x_0)| + \dots + \log|f'(x_n)|] / n$$

e se la quantità $[\log|f'(x_0)| + \dots + \log|f'(x_n)|] / n$ (la media aritmetica dei logaritmi delle derivate che si incontrano lungo la traiettoria) converge, quando n diventa molto grande verso un numero λ allora $f^{medico} \cong 2^\lambda$ equindi

$$E(n) \cong E(0) * 2^{\lambda n}$$

il numero che conta allora è la **media dei logaritmi** delle derivate lungo la traiettoria.

Questo numero λ viene chiamato **Coefficiente di Lyapunov della funzione f nel punto x_0** .

Più grande è λ e più grande è la velocità (media) di allontanamento di traiettorie che partono vicine. λ può anche essere negativo, in quel caso l'errore iniziale invece di aumentare diminuirà e in questo caso traiettorie che partono inizialmente vicine a x_0 si avvicineranno sempre più alla traiettoria di x_0 .

Per esempio il sistema dato da $f(x) = (1/2) * x$ ha $\lambda = -1$ per tutti i punti e infatti ogni condizione iniziale da luogo ad una traiettoria che converge molto velocemente allo 0 (infatti se la condizione iniziale è x_0 si ha $x_1 = x_0/2, \dots, x_n = x_0/(2^n)$).

Ricapitolando

Dato un sistema dinamico ed una condizione iniziale esiste un modo di ottenere un numero che viene chiamato coefficiente di Lyapunov ... e ...

Coeff. di Lyapunov > 0 implica:

- *Orbite che partono vicine si allontanano
- *Errori si amplificano
- *Il sistema non è stabile se leggermente perturbato
- *Chaos

Coeff. di Lyapunov < 0 implica:



- *Traiettorie che partono vicine restano vicine
- *Comportamento "prevedibile"
- *Stabilita'

E per $\lambda = 0$? Qualitativamente può succedere di tutto, le orbite si possono allontanare, ma non a velocità esponenziale.

Lettore che hai resistito fino a questo punto rilassati, le formule sono quasi finite.

L'informazione

Abbiamo detto che per noi un comportamento complesso è un comportamento la cui evoluzione richiede (al passare del tempo) molta informazione per essere descritta. In questo paragrafo si chiarisce cosa si intende per quantità di informazione e come la misuriamo.

Ci sono molti approcci possibili alla definizione del concetto di informazione e quantità di informazione contenuta in una sequenza di caratteri (che in linguaggio informatico si chiama anche stringa). Noi vorremmo avere una nozione che ci permetta, data una singola stringa S , di ottenere un numero $I(S)$ che considereremo come 'la quantità di informazione contenuta in S '. La misura della quantità di informazione contenuta in S quindi dipenderà solo da S e non da altre informazioni sul contesto in cui appare la stringa.

Come molti sanno, in ogni computer esistono dei programmi che comprimono i files. Questi programmi sono chiamati in gergo 'zippatori' (Winzip, gzip, Bzip2 e molti altri).

*** Cos'è di preciso uno zippatore?** Lo zippatore prende un file grosso e spesso lo codifica in un file più piccolo, che può essere archiviato risparmiando spazio sul disco del computer. Il file compresso poi all'occorrenza può essere decodificato per riottenere il file originale, così come era. Senza perdere informazioni.

La maggior parte dei documenti di testo (scritti in italiano ad esempio) può essere compressa ad un terzo della dimensione iniziale, ad esempio un file lungo 30K viene archiviato in un file compresso di 10K.

Kolmogorov. Molto velocemente accenneremo a cosa si tratta senza entrare nei dettagli.

La complessità di Kolmogorov di una stringa è la lunghezza del minimo programma (che eseguito da un certo computer) restituisce come output la stringa. In un certo senso l'associazione stringa  minimo programma può essere pensata come una compressione ottimale della stringa. Purtroppo questa associazione non può essere fatta da nessun algoritmo (da nessuna procedura effettiva finita).

Più precisamente si può dimostrare che non può esistere nessun programma (scritto in un qualsiasi linguaggio di programmazione) che calcoli la complessità di Kolmogorov delle stringhe.

Quindi la complessità di Kolmogorov è uno strumento teorico molto potente (che fra l'altro consente di dimostrare in modo semplice teoremi come il famoso teorema di incompletezza di Gödel) ma non è utilizzabile in pratica.

Per una esposizione divulgativa molto divertente di questa teoria clicca qui [Chiacchierata sulla teoria algoritmica dell'informazione \(V.Benci\)](#)

Complessità e rapporto di compressione

Adesso vediamo come si applica la quantità di informazione per misurare il grado di caoticità-complexità di un sistema.

Abbiamo un sistema dinamico che genera una traiettoria x_1, x_2, \dots, x_n che è una successione di punti dello spazio.

Adesso dividiamo lo spazio in 2 regioni: **A** e **B** (attenzione, per i semplici sistemi che considereremo bastano due regioni, ma in generale bisogna considerare anche più di due regioni), e associamo alla traiettoria x_1, x_2, \dots, x_n la stringa ottenuta sostituendo ad ogni punto della traiettoria l'insieme (**A** o **B**) che lo contiene, in questo modo si ottiene da una traiettoria di n punti una stringa binaria di n caratteri.

Ad esempio se si considera la mappa logistica come sistema dinamico, lo spazio è $[0,1]$, una buona scelta dei due insiemi è data da $A=[0,1/2)$, $B=[1/2,1]$ la successione ottenuta prima $x_0=0.1$
 $x_1=0.36$, $x_2=0.64$, $x_3=0.9216$, $x_4=0.289013\dots$, $x_5=0.82193\dots$ si trasforma nella seguente successione **AABBAB...**

Quindi ad una traiettoria di lunghezza n è associata una stringa binaria $s(n)$.

Adesso, come discusso prima misuriamo l'informazione contenuta in $s(n)$ utilizzando il nostro algoritmo di compressione Z , comprimiamo $s(n)$, ottenendo una stringa $Z(s(n))$ di cui misuriamo la lunghezza e questa sarà la quantità di informazione contenuta in $s(n)$.

Riassumendo Indichiamo con I la funzione che misura la quantità di informazione contenuta nelle stringhe e quindi definiamo

$$I(s(n)) = \text{lunghezza}(Z(s(n))).$$

Come abbiamo visto, da un sistema dinamico si possono ottenere traiettorie lunghe quanto si vuole e

quindi stringhe associate lunghe quanto si vuole.

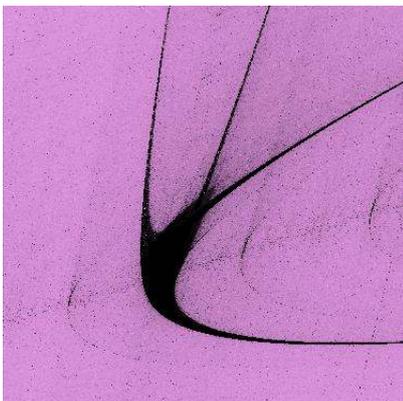
Se il sistema è caotico la stringa non sarà descritta da regole semplici e quindi conterrà molta informazione (rispetto alla sua lunghezza), mentre la stringa associata ad un sistema prevedibile (ad esempio un sistema periodico) conterrà poca informazione.

Quello che vogliamo fare adesso è definire un criterio che individui le stringhe più caotiche, siccome abbiamo a che fare con stringhe di varia lunghezza consideriamo la quantità media di informazione contenuta in un carattere della stringa.

Quindi considereremo la seguente quantità $K(s(n)) = I(s(n))/n$ (con n molto grande).

$K(s(n))$ sarà anche chiamato *rapporto di compressione della stringa $s(n)$* (infatti è il rapporto fra la lunghezza della stringa originale e quella della stringa compressa). Questo rapporto sarà un numero alto se la stringa contiene molta informazione e viceversa sarà un numero basso se la stringa contiene poca informazione e quindi si comprime molto rispetto alla sua lunghezza, questo indica che lo zippatore ha trovato delle forti regolarità nella stringa ed usa queste regolarità per comprimerla al meglio.

Come abbiamo fatto le figure



Vediamo come si ottengono le figure: abbiamo detto che un sistema dinamico è una coppia (spazio, funzione) la quale agisce sullo spazio e genera una dinamica. Supponiamo che da ora in poi lo spazio che ci interessa sia l'intervallo $[0,1]$ e consideriamo diverse funzioni, che daranno luogo a diversi tipi di dinamica.

Prima ad esempio era stata considerata la funzione $f(x) = 4x(1-x)$ sostituendo un altro numero al posto del 4 si possono avere tipi di dinamica completamente diversi. Cioè, se si considera la famiglia di funzioni $f(x, p) = 4p x(1-x)$ al variare di p fra 0 e 1 si possono avere diversi tipi di dinamica, da una dinamica *periodica* (quando p è piccolo) che ha coefficiente di

Lyapunov negativo in tutti i punti ad una dinamica caotica (quando p è grande) con coefficiente positivo.

Questa famiglia ad un parametro di sistemi dinamici è molto famosa e studiata (mappe logistiche). Per qualche ulteriore spiegazione rimandiamo al seguente link sulla [Mappa logistica](#).

Nel nostro caso si considera una famiglia a due parametri di funzioni dall'intervallo in se stesso $f(x,p,q)$ date dalla formula

$$f(x,p,q) = P(p, Q(q,x))$$

dove

$$P(x) = 4p(x(1-x))$$

e

$$Q(x) = 4q(x(1-x))$$

Come si vede quindi ogni funzione $f(x,p,q)$ è data dalla composizione di due funzioni logistiche con parametri anche diversi. I parametri possono ambedue variare fra 0 e 1 . Ogni possibile coppia di parametri (p,q) (che variano fra 0 e 1) può essere considerata come un punto di un quadrato. Ogni punto del quadrato che contiene la figura quindi corrisponde ad una funzione. Si calcola il coefficiente di Lyapunov di questa funzione (rispetto a qualche punto iniziale x scelto a caso, grosso modo in questi casi il coefficiente non dipende dal punto) e si colora il punto del quadrato in base al numero che si ottiene.

Il risultato sono quegli strani frattali pieni di tentacoli che si vedono nelle figure. Il primo a disegnare questo tipo di figure è stato Markus (Scientific American, sept 1991) per questo motivo in genere questo tipo di frattali vengono chiamati *frattali di Markus-Lyapunov*.

Per disegnare le figure che rappresentano la complessità si ripete la stessa costruzione considerando le stesse funzioni di prima si calcola il rapporto di compressione invece del coefficiente di Lyapunov.

Quindi ricapitolando per disegnare le figure:

Frattali di Markus-Lyapunov

consideriamo un punto (p,q) del quadrato, a questo punto corrisponderà una funzione $f(x,p,q)$.

Si considera $f(x,p,q)$, si sceglie un punto x a caso nel nostro spazio $[0,1]$ e si calcola il coefficiente di Lyapunov della funzione $f(x,p,q)$ nel punto x .

In base al valore che si ottiene si colora il punto del quadrato.



I punti con coefficiente di Lyapunov basso saranno colorati con colori scuri, mentre i punti che corrispondono a sistemi dinamici aventi coefficiente di Lyapunov più alto saranno colorati con colori più chiari.

Frattali della complessità

Per disegnare la figura:

consideriamo un punto (p, q) del quadrato, a questo punto corrisponderà una funzione $f(x, p, q)$.

Si considera $f(x, p, q)$, si sceglie un punto x a caso nel nostro spazio $[0, 1]$ si sceglie una partizione che genera una stringa simbolica e si calcola il rapporto di compressione della mappa $f(x, p, q)$ nel punto x .

In base al valore che si ottiene si colora il punto del quadrato.

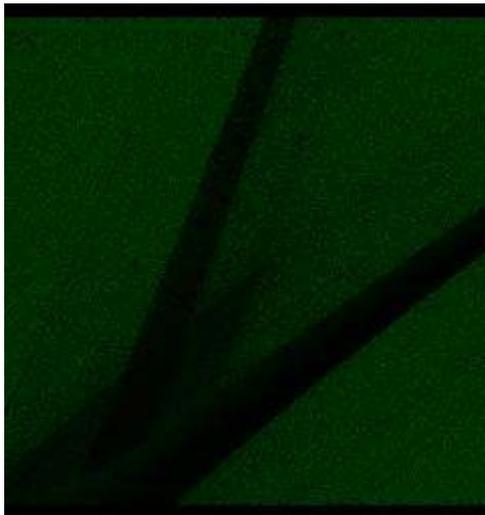


Figura qui sopra

I punti con complessità (rapporto di compressione) alta saranno colorati con colori chiari mentre quelli con complessità bassa saranno colorati con colori più scuri. Ad esempio nella figura qui sopra i punti con rapporto di compressione molto vicino a 0 sono neri.

Stefano Galatolo 2002-04-29