Alla ricerca della complessità

esame di caos, complessità e informazione

Marcello Seri

marcello.seri@studenti.unicam.it

Università degli Studi di Camerino

Definizioni

Innanzitutto dobbiamo precisare cosa intendiamo per sistema dinamico.

Da un punto di vista topologico, un **Sistema Dinamico** è una coppia (X, f) in cui X è uno *spazio metrico (generalmente compatto)* e $f : X \to X$ è una *trasformazione continua*.

Ci sono varie tecniche che ci permettono di tentare la comprensione locale o globale delle dinamiche del sistema. Ognuna di queste necessita di una modellizzazione appropriata del processo che dobbiamo studiare.

Per il momento lo strumento di cui ci serviremo è la cosiddetta **Dinamica Simbolica**.

Definizioni

L'obiettivo della *Dinamica Simbolica* è guardare ciò che accade non punto per punto, ma osservando lo spazio come se fosse una scacchiera discreta.

Partizioniamo X ed associamo un simbolo ad ogni elemento della partizione. Ad esempio:
sia $\mathcal{P} = \{P_0, P_1, \ldots, P_n\}$ una partizione di X cioè

$$X = \bigcup_{i=0}^{n} P_i \text{ tali che } P_i \cap P_j = \emptyset \text{ se } i \neq j$$

e sia $\mathcal{A} = \{0, 1, \ldots, n\}$ un alfabeto (abbiamo considerato quello formato dagli indici degli elementi delle partizioni).

Questa introduzione sarebbe inutile se non fosse possibile instaurare una relazione tra i punti dello spazio, gli elementi che abbiomo introdotto e le dinamiche del sistema.
Andiamo a vedere se e quando questo sia possibile...

Operiamo sulla partizione

Possiamo definire una mappa $\pi : X \to \mathcal{A}^{\mathbb{N}}$ come segue:

$$\pi(x) = \omega_0\omega_1\omega_2\dots \quad \text{tali che} \quad f^k(x) \in P_{\omega_k} \quad \text{con} \quad k \geqslant 0$$

ossia etichettiamo l'iterata k-esima con l'indice dell'area a cui appartiene il punto.

In altre parole, se supponiamo che $\omega_0\omega_1\omega_2\dots\omega_{k-1}$ sia una parola ammissibile, essa rappresenta il tratto di storia del sistema ottenuto iterando k volte la f sulla x.

Sorgo spontanei alcuni dubbi:

- come stabiliamo se la partizione mi permette di rappresentare fedelmente la dinamica del sistema?
- cosa succede cambiando partizione?

La dinamica è ben definita?

Cerchiamo di capire se data una sequenza infinita $\omega_0\omega_1\omega_2\ldots\omega_k\ldots$ è possibile risalire univocamente alla x che l'ha generata...

Sapere che $f(x) \in P_{\omega_1}$ significa stringere lo spazio in cui si trova x in P_{ω_0}, infatti $f^{-1}(P_{\omega_1}) = \{y \in X : f(y) \in P_{\omega_1}\}$ e $f^{-1}(P_{\omega_1}) \cap P_{\omega_0} \neq \emptyset$ altrimenti $\nexists\, x$ con parola $\omega_0\omega_1\ldots$!

Sapere che $f^2(x) \in P_{\omega_2}$ da una ulteriore informazione su x, infatti $x \in f^{-2}(P_{\omega_0}) \cap f^{-1}(P_{\omega_1}) \cap P_{\omega_1}$ che è non vuoto per quanto appena mostrato.

Quindi assegnare $k+1$ simboli vuol dire che inizialmente il nostro sistema si trovava in $x \in \bigcap_{i=0}^{k} f^{-i} P_{\omega_i}$!

Questo non vuol dire, però, che passando al limite per $i \to \infty$ riusciamo ad ottenere un insieme di cardinalità 1.

Partizioni generanti

Quando accade che

$$\pi(x) = \pi(x') \text{ se e solo se } x = x'$$

o altrimenti che

$$\text{se } \pi(x) = \omega_0\omega_1\omega_2\ldots \text{ allora } x = \bigcap_{i \geqslant 0} f^{-i}(P_{\omega_i})$$

diciamo che la partizione $\mathcal{P}$ è **generante**.

Si trovano facilmente partizioni generanti ad esempio per mappe unimodali, per la mappa di bernoulli e per la mappa a tenda.

Parole ammissibili

Prima di cominciare a parlare della complessità restano solo altri due concetti da definire...

Consideriamo

$$f(x) = 2x \mod 1 : [0,1] \to [0,1]$$

la partizione $\mathcal{P} = \left\{ \left]0, \frac{1}{2}\right[, \left[\frac{1}{2}, 1\right[\right\}$ e l'alfabeto $\mathcal{A} = \{0, 1\}$

da tale dinamica possiamo ottenere *tutte* le parole di $\mathcal{A}^N$.

Parole non ammissibili

Se però considerassimo la mappa

$$f(x) = rx(1-x) : I_0 \cup I_1 \to I_0 \cup I_1$$

con r definito da $f^3(\frac{1}{2}) = \frac{r-1}{r}$, la partizione

$$\mathcal{P} = \left\{ I_0 = \left] f^2(\frac{1}{2}), \frac{1}{2} \right[, I_1 = \left[\frac{1}{2}, f(\frac{1}{2}) \right] \right\}$$

e l'alfabeto $\mathcal{A} = \{0, 1\}$

non potremmo ottenere le parole del tipo '$01^{2n}0$' dove '1^{2n}' indica '1' ripetuto $2n$ volte.

Definizione

Come abbiamo osservato può capitare che non tutte le parole possibili di una certa lunghezza possano essere generate dal sistema, ci possono cioè essere delle proibizioni.

Diremo che una parola $\omega_0\omega_1\ldots\omega_{N-1}$ è **ammissibile** per $(X, f, \mathcal{P})$ se

$$\bigcap_{k=0}^{N-1} f^{-k}(P_{\omega_k}) \neq \emptyset$$

In sostanza u è *ammissibile* se

$$\exists\, x \quad \text{tale che} \quad \begin{cases} x \in P_{\omega_0} \\ f(x) \in P_{\omega_1} \\ \vdots \\ f^N(x) \in P_{\omega_{N-1}} \end{cases}$$

Definizione

A questo punto, se

$$\mathcal{A}^* = \bigcup_{N \in \mathcal{N}} \mathcal{A}^N$$

è il *monoide libero di* $\mathcal{A}$,

definiamo **Linguaggio formale associato ad** $(X, f, \mathcal{P})$:

$$\mathcal{L}(X, f, \mathcal{P}) = \{u \in \mathcal{A}^* : u \text{ è ammissibile per } (X, f, \mathcal{P})\} \subseteq \mathcal{A}^*$$

Un primo approccio alla complessità

Supponiamo $\mathcal{A} = \{0, 1, \ldots, n\}$, uno delle prime cose che potremmo voler fare è contare il numero di parole di una certa lunghezza, in modo da avere una prima idea della ricchezza del linguaggio che stiamo considerando.

Definiamo a tal proposito la **funzione complessità**:

$$p(N) = \#\{u \in \mathcal{L} : |u| = N\}$$

In questo modo abbiamo due casi estremi con una enorme varietà di casi intermedi alcuni dei quali molto importanti... Quando la crescita è la più lenta possibile, è il caso dei *linguaggi Sturmiani*, la complessità è la più bassa possibile: $p(N) = N + 1$ (es. le rotazioni del cerchio modulo 1). Quando la crescita è la più veloce possibile, il linguaggio coincide con $\mathcal{A}^*$ e la complessità è esponenziale: $p(N) = (n + 1)^N$ (es. la mappa di Bernoulli).

Un primo approccio alla complessità

Se $p(N)$ cresce esponenzialmente (o meglio *la partizione generatrice si autoraffina proliferando esponenzialmente*) ha senso calcolare quale è il suo *tasso di crescita esponenziale*. Definiamo dunque l'**entropia della funzione** f **rispetto alla partizione** $\mathcal{P}$:

$$h(f, \mathcal{P}) = \lim_{N \to +\infty} \frac{\log p(N)}{N}$$

Se consideriamo l'estremo superiore di tale entropia otteniamo l'**entropia topologica della funzione** f

$$h(f) = \sup_{\mathcal{P}} h(f, \mathcal{P})$$

Per questaè rilevante il *teorema del Generatore* di Sinai che afferma che

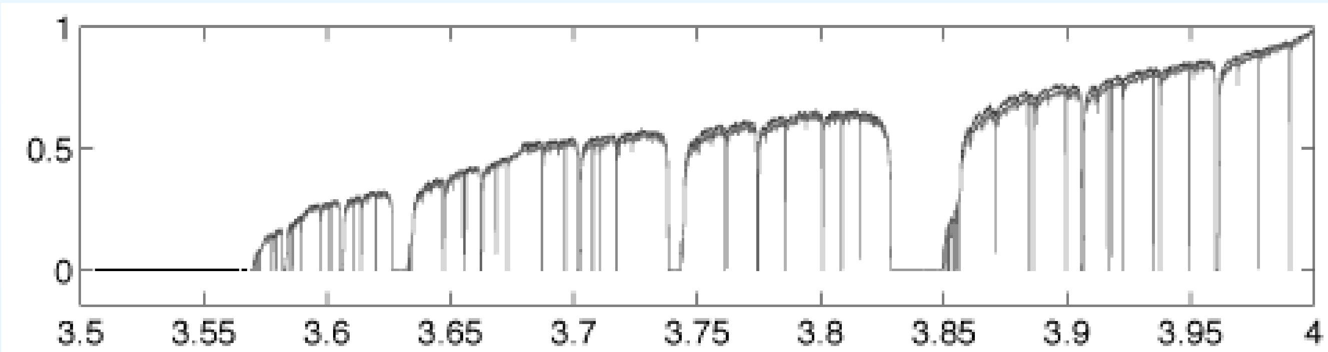
$$\mathcal{P} \text{ generante } \implies h(f, \mathcal{P}) = h(f).$$

Alcune considerazioni

In base a quanto osservato precedentemente, la mappa di Bernoulli ha entropia topologica $h(f) = \log_2 2 = 1$ (che è la massima possibile), mentre le rotazioni hanno entropia topologica $h(f) = 0$ (la minima).

Tra questi casi estremi ci sono moltissimi sistemi a crescita polinomiale (per cui con entropia nulla) che ancora non sono stati classificati e caratterizzati esaurientemente.

Chiameremo "*caotico*" (in senso lato) un sistema ad entropia positiva, diremo, invece, che è un sistema di "*caos debole*" se ha entropia nulla ed è non periodico.



Raffiniamo il concetto di entropia

Per quanto apra già scenari interessanti, il concetto di entropia che abbiamo definito è ancora rudimentale (si limita a contare). Cerchiamo di raffinarlo assegnando in qualche modo un peso alle parole.

Supponiamo dunque che (f, X) ammetta una *misura di probabilità invariante* μ, cioè tale che

$$\forall E \subset X \text{ misurablie, si ha } \mu(f^{-1}(E)) = \mu(E)$$

e costruiamo una distribuzione di probabilità sulle parole del linguaggio di lunghezza N

$$p(\omega_0\omega_1 \ldots \omega_{N-1}) = \mu\left(\bigcap_{k=0}^{N-1} f^{-k}(P_{\omega_k})\right)$$

L'entropia metrica

Per analogia con la prima definizione di entropia che abbiamo dato, possiamo definire prima una sorta di contatore delle parole che tenga conto del peso

$$H(\mathcal{P}, N) = -\sum_{|u|=N} p(u) \log p(u)$$

e, a questo punto, l'**entropia della mappa** f **sulla partizione** $\mathcal{P}$ **relativa alla misura** μ

$$h_\mu(f, \mathcal{P}) = \lim_{N \to +\infty} \frac{H(\mathcal{P}, N)}{N}$$

da cui si ottiene l'**entropia di Kolmogorov-Sinai**

$$h_\mu(f) = \sup_{\mathcal{P}} h_\mu(f, \mathcal{P})$$

Anche in questo caso vale il Teorema del generatore di Sinai.

Legame tra le entropie

Abbiamo definito una nuova entropia che sembra essere più fine della precedente entropia topologica perchè ci permette di pesare le parole che andiamo a considerare. Ma cè qualche relazione che lega queste due definizioni?

Sì, in primo luogo sono entrambe caratterizzazioni per il sistema dinamico (X, f) che danno una misura complessiva della ricchezza e della complessità del linguaggio che essa genera, inoltre indicano entrambe se il sistema produce informazione nella sua evoluzione, ma soprattutto vale il seguente *principio variazionale*

$$h(f) = \sup_{\mu} h_{\mu}(f)$$

(quando esiste μ che soddisfa tale uguaglianza, viene chiamato *misura di massima entropia*).

Considerazioni sulla Complessità

Giunti a questo punto è doveroso fermarsi per rifletere su quanto fatto.

Volevamo trovare *la* definizione di Complessità ed in qualche modo l'abbiamo fatto... ma che vuol dire complessità?

Il problema non è banale ed è tuttora senza una risposta precisa (per lo meno per quanto ne so io)... in altre parole dipende dal punto di vista, dalla sfumatura di significato che ci interessa e da molti altri fattori a volte più soggettvi che oggettivi!

La parzialità dell'approccio che abbiamo seguito è evidente, abbiamo cercato la complessità come indicatore della ricchezza di lignuaggi che stabiliamo a priori, come una sorta di indice che rappresenta la velocità con cui cresce il vocabolario che possiamo generare.

Considerazioni sulla Complessità

Inoltre per calcolare la complessità ricorriamo a delle classi di parole e ne cerchiamo il comportamento al limite.

Se conoscessimo una parola (infinita) e volessimo sapere qualche cosa sulla sua complessità? Abbiamo speranze di farlo?

C'è un lavoro di Kolmogorov che ha mostrato che sfruttando il concetto di Macchina di Turing universale è possibile definire la complessità secondo questo punto di vista.

Kolmogorov definisce la **complessità algoritmica**, in sostanza una misura della difficoltà di riprodurre la parola che consideriamo.

Un breve cenno sulla complessità algoritmica

Kolmogorov ha dimostrato che esiste una MdT universale che produce una qualunque parola u che ha lunghezza minima.

Consideriamo $(X, f, \mathcal{P})$, prendiamo $x \in X$ e costruiamo $u_N(x) = \omega_0\omega_1 \dots \omega_{N-1}$ e chiamiamo $K(v)$ la *lunghezza del programma più breve* (scritto in binario) *che descrive una MdT universale che produce come output la parola* v.

Definiamo **complessità algoritmica della traiettoria di x sulla partizione** $\mathcal{P}$ la quantità:

$$K(x, f, \mathcal{P}) = \lim_{N \to +\infty} \frac{K(u_N(x))}{N}$$

E analogamente a quanto fatto fino ad ora definiamo **complessità di x** il valore $K(x, f) = \sup_{\mathcal{P}} K(x, f, \mathcal{P})$

Legame tra le entropie

Strano ma vero, questa definizione di complesità è strettamente legata all’entropia topologica, infatti $K(x, f) \leqslant h(f)$ e vale il *Teorema di Brudno* per cui

Se μ è una misura f-invariante ed ergodica
allora $K(x, f) = h(f)$ μ-q.o.

A titolo informativo, μ si dice *ergodica* se $\forall\ E \subset X$ misurabile ed f-invariante ($E = f(E)$) si ha $\mu(E) = 0$ oppure $\mu(E) = 1$.

Considerazioni sulla Complessità

In questo modo abbiamo risolto il problema delle parole di cui non conosciamo il linguaggio generatore, ma resta il fatto (che avevo volutamente tralasciato) che calcolando una qualunque delle entropie troviamo ai valori estremi la situazione completamente aleatoria e quella eccessivamente regolare, mentre lasciamo in mezzo tutto quello che è ricco e dettagliato... se ci pensiamo bene, intuitivamente vorremmo la complessità massima al centro e minima agli estremi, una sorta di gaussiana.

In sintesi con l'entropia abbiamo una misura dell'aleatorietà.

È legittimo, allora, chiedersi se possono esserci altri tipi di definizioni di complessità o se si possano osservare altri suoi aspetti...

Un cenno alla gerarchia di Chomski

Un esempio di approccio diverso è la *gerarchia di Chomski* che tenta di classificare i linguaggi formali in base alla complessità della loro grammatica definendo 4 classi:

1. **Regular**, rappresentabili mediante *automi finiti*
2. **Context-free**, rappresentabili mediante *automi push-down*
3. **Context-sensible**, rappresentabili mediante *automi linear-bounded*
4. **Unrestricted**, rappresentabili mediante *macchine di Turing*

Tale gerarchia non è esaustiva, tra le ultime due classi ad esempio ci sono i *linguaggi generati da sostituzioni* che sono un insieme importante.

Resta comunque un ottimo esempio di come ci siano stati (ci sono e ci saranno) altri approcci alla complessità, ognuno coi suoi pregi e i suoi difetti, ognuno legato a degli aspetti particolari e chissà magari un giorno anche uno completo ed esaustivo...