

Soluzione Compito di Modelli Biologici Discreti

Docente: Dr Giuditta Franco

appello 3 febbraio 2012

1 I parte di programma

1. **Definire cos'è una dinamica. Indicare una particolare dinamica, di cui si è vista un'analisi qualitativa, e delineare i punti principali di tale analisi.**

Def (dinamica). Una funzione da un insieme di istanti temporali ad un insieme di stati (detto spazio delle fasi) di un sistema.

Nel corso si è affrontato uno studio qualitativo di dinamiche di cui non abbiamo la funzione esplicita, ma soltanto la forma ricorrente o differenziale. Per esempio, la mappa logistica, o il modello preda-predatore di Lotka-Volterra.

In un'analisi qualitativa, di una dinamica si studiano i comportamenti asintotici e oscillanti, ovvero si individuano i punti fissi, la loro iperbolicità, e le orbite periodiche (con corrispondente lunghezza o periodo di oscillazione).

2. **Analogamente a come studiato per la mappa logistica, trovare il punto fisso della seguente “tent map” $x(n+1) = \frac{1}{4}(1 - 2|x(n) - \frac{1}{2}|)$ e valutarne l'iperbolicità (ovvero dire se è attrattivo o repulsivo).**

Almeno tre possibili modi di procedere (intuitivo, logico-deduttivo, standard):

- (a) Come per la mappa logistica, nell'intervallo $[0,1]$ si consideri il grafico (a forma di “tenda” invece che di parabola) per la funzione che itera $f(x) = \frac{1}{4}(1 - 2|x - \frac{1}{2}|)$. Spostandosi sul grafico per valori consecutivi di $x(n)$ (o iterati di $f^n(x(0))$), tramite la bisettrice (del primo quadrante), che sovrasta il grafico, si visualizza la convergenza allo zero (punto fisso, attrattore della dinamica).
- (b) Si può osservare che, per qualsiasi valore arbitrario di $x(0) \in [0, 1]$, si ottiene $x(1) \leq \frac{1}{4}$, e quindi $x(n+1) = \frac{1}{2}x(n) \quad \forall n \geq 1$. Essendo infatti $|x(n) - \frac{1}{2}| \leq \frac{1}{2}$ per $0 \leq x(n) \leq 1$, si ha $2|x(n) - \frac{1}{2}| \leq 1$, e $\frac{1}{4}(1 - 2|x(n) - \frac{1}{2}|) \leq \frac{1}{4}$.
Si tratta dunque di una malthusiana (con coefficiente di crescita minore di 1) strettamente decrescente, che converge a zero (unico punto fisso, attrattore della dinamica).

- (c) Si studiano le due equazioni di ricorrenza: $x(n+1) = \begin{cases} \frac{1}{2}x(n) & x(n) \leq \frac{1}{2} \\ \frac{1}{2}(1 - x(n)) & x(n) \geq \frac{1}{2} \end{cases}$

Imponendo $x(n+1) = x(n)$ si trova la soluzione nulla per la prima equazione, e $x(n) = \frac{1}{3}$ per la seconda, che non appartenendo al dominio di definizione non è soluzione valida. L'unico punto fisso della dinamica, lo zero, risulta attrattore (secondo un teorema noto) perché $|f'(0)| = \frac{1}{2} < 1$.

3. **Spiegare il modello “a ragnatela” della domanda-offerta, specificando perché si chiama così.**

Si vuole individuare la funzione $p(k)$, prezzo di un bene nel tempo, affinché la domanda $d(k)$ del bene incontri l'offerta $s(k)$.

Sia $d(k) = d_0 - ap(k)$, con d_0 valore iniziale della domanda, e $a > 0$ influenza positiva del prezzo sulla domanda, e $s(k+1) = s_0 - bp(k)$, dove s_0 è il valore iniziale dell'offerta e $b > 0$ misura l'influenza (in negativo) del prezzo sulla produzione futura dell'offerta. Si tratta di risolvere l'equazione di ricorrenza (lineare, non omogenea): $ap(k+1) + bp(k) = d_0 - s_0$, con soluzione particolare $\frac{d_0 - s_0}{a+b}$ e soluzione dell'omogenea associata $p(k) = A(-\frac{b}{a})^k$, dove $A = p(0) - \frac{d_0 - s_0}{a+b}$.

Il prezzo $p(k) = (p(0) - \frac{d_0 - s_0}{a+b})(-\frac{b}{a})^k + \frac{d_0 - s_0}{a+b}$ converge al punto fisso $\frac{d_0 - s_0}{a+b}$ per $b < a$ (ovvero quando l'offerta è meno sensibile al prezzo della domanda), oscilla per $a = b$, e diverge altrimenti. Collegando passi consecutivi della funzione iterata del prezzo sulle rette della domanda e dell'offerta si disegna una ragnatela, da cui il nome del modello.

4. **Sia data una popolazione che evolve annualmente secondo l'equazione malthusiana $x(n+1) = 1.5 \cdot x(n) + m$, con m emigrazione annua costante. Partendo da una popolazione iniziale $x(0)$ di 10 milioni di individui, si chiede quale sarà il destino della popolazione (comportamento dinamico nel futuro: esplosione, estinzione, assetamento su un certo valore, maggiore o minore di quello iniziale), nei seguenti due casi:**

- (a) **$m = -4$ milioni?** Essendo $10 > \frac{4}{0.5}$, la popolazione asintoticamente esplose
 (b) **$m = -5.5$ milioni?** Essendo $10 < \frac{5.5}{0.5}$, la popolazione si va ad estinguere per l'alto flusso migratorio in uscita, nonostante il fattore di crescita maggiore di 1.

2 II parte di programma

1. **Sia dato il seguente insieme di riscrittura:** $\left\{ \begin{array}{l} c \rightarrow a \\ a \rightarrow b \\ a \rightarrow bc \\ b \rightarrow bc \\ b \rightarrow ab \end{array} \right.$ **di cui indichiamo con**

r_1, r_2, r_3, r_4, r_5 le regole nell'ordine in cui compaiono, e sia $M = abccab$ un multinsieme di partenza.

- (a) **Indicare quali e quanti multinsiemi si ottengono applicando ad M una volta le regole date, secondo una strategia di massimo parallelismo non deterministico.**

Si ottengono 9 multinsiemi $\{a^4b^4, a^4b^4c, a^4b^4c^2, a^3b^4c, a^3b^4c^2, a^3b^4c^3, a^2b^4c^2, a^2b^4c^3, a^2b^4c^4\}$.

- (b) **Le regole sopra si riferiscano ora ad un sistema metabolico, di cui M rappresenta lo stato iniziale $X(0)$. Calcolare lo stato $X(1)$, secondo le seguenti funzioni di flusso: $u_1 = f_1(a, b, c) = ab, u_2 = f_2(a, b, c) = c^2, u_3 = f_3(a, b, c) = 2a, u_4 = f_4(a, b, c) = a, u_5 = f_5(a, b, c) = c$.**

Secondo EMA: $X[1] = A \times U[0] + X[0]$, ove, nel nostro caso $A = \begin{pmatrix} 1 & -1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 \end{pmatrix}$,

$$U[0] = \begin{pmatrix} 4 \\ 4 \\ 4 \\ 2 \\ 2 \end{pmatrix}, \text{ e } X(0) = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, \text{ quindi } X(1) = \begin{pmatrix} 0 \\ 10 \\ 4 \end{pmatrix}.$$

2. Si diano le espressioni regolari associate ai linguaggi riconosciuti dai seguenti due automi.

$$a(bca)^*a \quad \left\{ \begin{array}{ll} q_0 a \rightarrow q_1 & q_0 \text{ (iniziale)} \\ q_1 b \rightarrow q_2 & \\ q_2 c \rightarrow q_0 & \\ q_1 a \rightarrow q_3 & q_3 \text{ (finale)} \end{array} \right.$$

$$(aa + bb)a^*b \quad \left\{ \begin{array}{ll} q_0 a \rightarrow q_1 & q_0 \text{ (iniziale)} \\ q_0 b \rightarrow q_2 & \\ q_1 a \rightarrow q_3 & \\ q_2 b \rightarrow q_3 & \\ q_3 a \rightarrow q_3 & \\ q_3 b \rightarrow q_4 & q_4 \text{ (finale)} \end{array} \right.$$

3. Riportare l'algoritmo DNA di Sakamoto per risolvere SAT, sia in termini formali che di implementazione.

Discuterne la correttezza, la completezza, l'efficienza, e la biocomplessità (nello spazio e nel tempo). Indicarne un punto di forza e uno di debolezza.

Un'istanza 3-SAT(n, m) data viene risolta prima generando le molecole DNA che contengono uno dei tre letterali per ogni clausola (notazione: $C_i = l_1^{(i)} \vee l_2^{(i)} \vee l_3^{(i)}$) e poi estraendo i corrispondenti assegnamenti (non contraddittori) che rappresentano le soluzioni.

INPUT: $P = \{\alpha C_1, \overline{C_1} L_1 C_2, \dots, \overline{C_{m-1}} L_{m-1} C_m, \overline{C_m} L_m \beta \mid |C_i| = |\overline{C_i}| = 20 \text{ sticky ends}, |\alpha| = |\beta| = 20, L_i \in \{l_1^{(i)}, l_2^{(i)}, l_3^{(i)}\}, |L_i| = 20, \gamma_z \in L_i, i = 1, \dots, m, \gamma_z \text{ sito di restrizione dell'enzima Enz}\}$.

ALGORITMO:

- (a) $P := Lig(C(P));$ \ \ Assemblaggio dei 3^m assegnamenti soddisfacenti le m clausole
- (b) $P = PCR(\alpha, \overline{\beta})(P);$ \ \ amplificazione del materiale
- (c) $P = H(P);$ \ \ riscaldamento del pool ad elevata temperatura, per garantire che tutti i filamenti diventino singoli
- (d) $P = Enz(C(P));$ \ \ abbassamento brusco della temperatura, per garantire che la formazione di hairpin preceda la rinaturazione, e taglio immediato dell'enzima
- (e) $P = El_{40(m+1)}(P);$ \ \ selezione dei filamenti (ancora singoli) rimasti della lunghezza originaria
- (f) $P = PCR(\alpha, \overline{\beta})(P);$ \ \ amplificazione del materiale
- (g) if $El(P) \neq \emptyset$ then X= Yes else X=NO

OUTPUT: Value of X.

L'algoritmo è corretto, perchè se nel pool finale si trova qualche molecola, quella ha superato integra la fase di taglio dell'enzima nel passo (d), e quindi codifica un assegnamento coerente che verifica almeno un letterale in ognuna delle clausole. Non è completo, per via dell'inefficienza del passo (d), ove le tempistiche di abbassamento della temperatura sono cruciali quanto delicate, e qualcuna delle molecole soluzione potrebbe rinaturarsi e venire erroneamente tagliata dall'enzima (che opera su ogni occorrenza in doppio filamento di γ_2): in questo caso non tutte le soluzioni verrebbero trovate. L'algoritmo ha biocomplexità nello spazio esponenziale (3^m) rispetto al numero di clausole, e costante nel tempo.

Punti di forza dell'algoritmo si individuano nella complessità costante nel tempo e nell'efficienza del tipo di operazioni utilizzate, mentre punti di debolezza si trovano nella complessità eccessiva nello spazio e nella delicatezza di esecuzione del passo (d).

4. Indicare l'output del seguente bioalgoritmo, ricordando che $XPCR_\gamma$ è la procedura che implementa su un pool di molecole DNA la regola di riscrittura di "γ-splicing":

INPUT: Pool di sequenze tutte note $P = \{\alpha X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 \beta, \alpha X_1 Y_2 X_3 Y_4 X_5 Y_6 X_7 Y_8 \beta, \alpha Y_1 X_2 Y_3 X_4 Y_5 X_6 Y_7 X_8 \beta, \alpha Y_1 Y_2 Y_3 Y_4 Y_5 Y_6 Y_7 Y_8 \beta\}$ tali che $|X_i| = |Y_i| = 20 \quad \forall i = 1, \dots, 8$.

Per esempio, se X_i e Y_i sono sequenze DNA che codificano i due valori di verità $\{0, 1\}$ di una (su otto) variabile booleana x_i , nel pool troviamo i quattro assegnamenti 11111111, 00000000, 10101010, e 01010101, con molecole aventi tutte un prefisso α e un suffisso β .

ALGORITMO: for $i=2,7$

$(P_1, P_2) := split(P)$;

$P_1 := XPCR_{X_i}(P_1)$ and $P_2 := XPCR_{Y_i}(P_2)$;

$P := mix(P_1, P_2)$;

end

OUTPUT: $P = \{\alpha Z_1 Z_2 Z_3 Z_4 Z_5 Z_6 Z_7 Z_8 \beta \mid Z_i = \{X_i, Y_i\}, \quad |X_i| = |Y_i| = 20 \quad \forall i = 1, \dots, 8\}$.