

Modello SIR differenziale e automa cellulare

Simone Rossi Tisbeni

August 16, 2018

Abstract

In questo paper viene studiato il Modello SIR per la diffusione delle epidemie e sono confrontati due diversi tipi di modello: il classico sistema di equazioni differenziali ordinarie, e un modello a Cellular Automata che tiene conto della spazialità del sistema. I due modelli vengono confrontati per valutare la bontà dei risultati e le possibili condizioni di applicazione. Una rassegna sullo stato dell'arte introduce l'elaborato.

Introduzione

Lo studio delle epidemie, come distribuzione e frequenza di malattie ed eventi di rilevanza sanitaria nella popolazione, è di grande interesse per il settore biomedico.

Avvalendosi di statistica e modelli matematici è possibile studiare la storia naturale della malattia e programmare piani di controllo per monitorarla e limitarne la diffusione tra la popolazione.

Il metodo più diffuso per lo studio di malattie infettive sono i cosiddetti modelli compartimentali, in cui la popolazione interessata è divisa in classi, con caratteristiche uguali tra gli individui di una stessa classe.

Il modello SIR è una delle forme più semplici di modello compartimentale ed è stato proposto per la prima volta da W. O. Kermack e A. G. McKendrick [6] come caso particolare di quella che è nota come teoria di Kermack-McKendrick.

Il modello è perfetto per rappresentare epidemie senza ciclo vitale, in cui il *rate ri-*

produttivo R_0 determina lo scoppio o meno dell'epidemia [3] [4].

Questo tipo di modello presuppone l'interazione popolazione-popolazione non localizzata, in cui ogni individuo può entrare a contatto con l'intera popolazione.

Per studiare gli effetti della spazialità, viene confrontato il più classico modello differenziale con un modello basato su automi cellulari (CA), elaborato a partire dagli articoli di White et al. [8] e Bouaine e Rachik [1].

Il concetto di automa cellulare è stato introdotto da J.VonNeumann e S.Ulman [7] e reso noto da J.Conway con il famoso *game of life* [2].

Gran parte degli studi sui CA sono però dovuti a S. Wolfram negli anni '80 che nel suo controverso libro *A new kind of Science* [9] introduce i CA a 2 dimensioni e la teoria per cui gli automi cellulari siano macchine di Turing universali, ovvero ogni sistema fisico reale computabile può essere modellizzato da un'appropriata implementazione di un CA.

In questo paper verranno presentati e

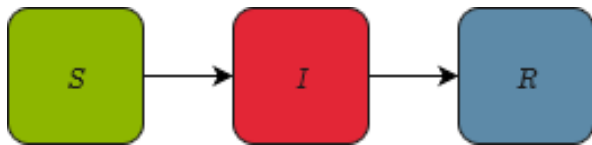


Figura 1: Diagramma di flusso del modello SIR

analizzati due differenti modelli.

Nella prima sezione, verrà descritto il modello SIR così come presentato da Kermack-McKendrick, introducendo in maniera qualitativa il rate riproduttivo.

Nella seconda sezione viene proposto il classico modello differenziale, viene data una definizione analitica di rate riproduttivo e confrontati due differenti metodi di integrazione del sistema.

Nella terza sezione viene introdotto il concetto di automa cellulare, viene presentato il modello proposto e derivato il concetto di rate riproduttivo dalle regole di transizione. Infine, viene analizzato l'andamento spaziale del modello in funzione degli indici di vicinato utilizzati.

L'ultima sezione è dedicata al confronto tra i modelli, con una breve discussione sul possibile campo di applicazione.

In appendice al paper sono presentati gli script in pseudocodice utilizzati per l'integrazione del modello e la simulazione su reticolo CA.

1 Modello SIR

Il modello prevede la divisione della popolazione in tre classi indipendenti: S, che indica la frazione di popolazione suscettibile alla malattia; I, che indica la frazione di popolazione infettata dalla malattia ed infettiva; R, che indica la frazione di popolazione rimossa, a causa di guarigione, immunizzazione o decesso, che è quindi incapace di trasmettere la malattia. Il modello è adatto a simulare malattie

che si trasmettono tramite contatto tra individui e che conferiscono immunità in seguito alla guarigione, con un'espansione del morbo sufficientemente rapida da rendere trascurabili le morti e le nascite naturali degli individui.

Assumendo che nessun individuo venga aggiunto al gruppo dei suscettibili, ignorando quindi eventuali nascite o immigrazioni, e che non vi siano morti naturali o emigrazioni, la variazione nel tempo del numero di individui suscettibili dipende esclusivamente dal numero di individui suscettibili stesso, dal numero di individui infetti e dalla probabilità che il contatto tra individui infetti e suscettibili porti al contagio.

Assumiamo, inoltre, che una frazione fissa di infetti passi al gruppo dei rimossi a ogni istante di tempo. Questa, più propriamente, rappresenta la frazione di individui che perde la capacità di infettare individui suscettibili.

In Figura 1 viene mostrato il diagramma che rappresenta il modello di transizione descritto.

Con queste assunzioni la popolazione totale del modello può essere considerata costante.

Un concetto fondamentale in epidemiologia è l'esistenza di termini soglia, valori critici di popolazione da superare perché avvenga lo scoppio di un'epidemia [3].

Parametro caratteristico per determinare i valori critici del modello è il *rate riproduttivo* R_0 , definibile come il numero medio di effettive infezioni da parte di un individuo infetto durante il periodo di infettività della malattia.

Intuitivamente si può vedere come il rate riproduttivo sia pari a 1 quando la malattia si trova in equilibrio endemico (cioè in uno stato stazionario positivo). Se mediamente un individuo infettivo trasmettesse la malattia a più di una persona, la malattia si diffonderebbe. Se la malattia venisse trasmessa mediamente a meno di una persona, si esaurirebbe nel tempo.

2 Modello differenziale

Tenendo conto delle definizioni della sezione precedente, e assumendo che la popolazione considerata abbia dimensione costante N sufficientemente grande da poter considerare la popolazione di ogni classe come una variabile continua invece che discreta, possiamo definire la dinamica temporale della popolazione delle differenti classi come:

$$\frac{dNS}{dt} = -\beta NS(t)I(t) \quad (1a)$$

$$\frac{dNI}{dt} = \beta NS(t)I(t) - \gamma NI(t) \quad (1b)$$

$$\frac{dNR}{dt} = \gamma NI(t) \quad (1c)$$

$$\text{con } R(t) + S(t) + I(t) = N \quad (2)$$

β è il *tasso di infettività*: il numero medio di trasmissioni per istante di tempo per individuo infetto. Quindi, βS è il numero medio di suscettibili infettati mediamente per ogni istante di tempo da un singolo infetto e βNIS è il numero medio di suscettibili infettati dalla classe di infetti per istante di tempo.

Il termine $-\gamma NI$ determina il ritmo di recupero della popolazione infetta. In particolare, $1/\gamma$ corrisponde alla durata media dell'infezione.

Se dividiamo ogni equazione per la popolazione N otteniamo la dinamica temporale delle frazioni di popolazione:

$$\frac{dS}{dt} = -\beta S(t)I(t) \quad (3a)$$

$$\frac{dI}{dt} = \beta S(t)I(t) - \gamma I(t) \quad (3b)$$

$$\frac{dR}{dt} = \gamma I(t) \quad (3c)$$

$$\text{con } S(t) + R(t) + I(t) = 1 \quad (4)$$

Ogni parametro nelle equazioni differenziali descritte è non-negativo e si considerano solo le soluzioni non-nulle, in quanto soluzioni

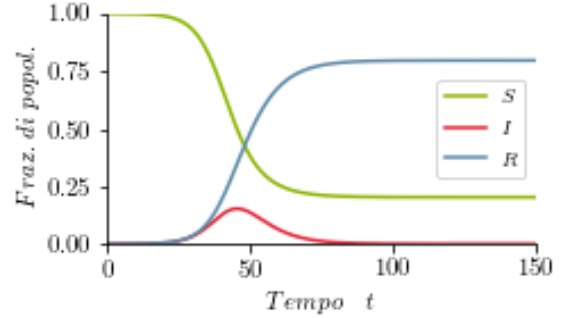


Figura 2: Curve di frazione di popolazione suscettibile, infetta e rimossa in funzione del tempo. I parametri sono $\beta = 0.4$ e $\gamma = 0.2$.

con frazioni di popolazione negative non hanno significato fisico.

Possiamo ricavare I da 3c e sostituire in 3a ottenendo

$$\frac{1}{S(t)} \frac{dS}{dt} = -\frac{\beta}{\gamma} \frac{dR}{dt} \quad (5)$$

e integrare:

$$S(t) = S_0 e^{-\beta R(t)/\gamma} \quad (6)$$

dove S_0 è la porzione di popolazione suscettibile al tempo $t = 0$.

Risolvendo per I eq. 4 e sostituendo con eq. 6 in eq. 3c otteniamo:

$$\frac{dR}{dt} = \gamma \left(1 - R(t) - S_0 e^{-\beta R(t)/\gamma} \right) \quad (7)$$

Se risolvessimo per R potremmo trovare S da eq. 6 e I da eq. 4, ma non esiste una soluzione esatta, il che rende necessari metodi numerici di integrazione, per ottenere un grafico analogo a quello di Figura 2.

Dal grafico si vede come la frazione di Suscettibili decresce monotonicamente al crescere del numero di infezioni, e la curva di popolazione rimossa cresce monotonicamente. La frazione di infetti invece aumenta allo scoppio

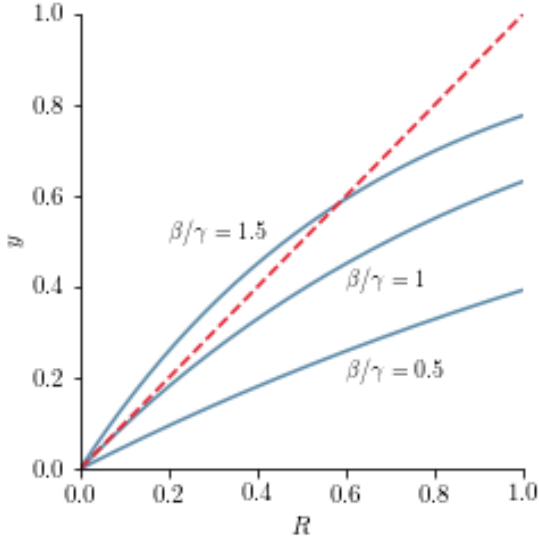


Figura 3: Andamento di $y(R) = 1 - e^{-\beta R/\gamma}$ per diversi valori di β/γ . La linea tratteggiata corrisponde a $y = R$

dell'epidemia, per poi decrescere al decorrere dell'infezione, tendendo a zero per $t \rightarrow \infty$.

Il valore dell'asintoto a cui tende R ha un importante significato in epidemiologia, in quanto corrisponde al numero di individui che hanno contratto l'infezione durante l'intero periodo dell'epidemia.

Può essere ottenuto da eq. 7 per $dR/dt = 0$, che risulta in $R(t) = 1 - S_0 e^{-\beta R(t)/\gamma}$

Il valore di S al tempo 0 può essere modellizzato in molti modi. Comunemente si assume che al tempo 0 il numero di individui infetti sia molto piccolo, e il resto della popolazione sia nella classe di Suscettibili:

$$S_0 = 1 - \frac{c}{N} \quad \text{con } c \ll N$$

Nel limite di una popolazione molto grande ($N \rightarrow \infty$) possiamo considerare $S_0 \simeq 1$ e ottenere

$$R(t) = 1 - e^{-\beta R(t)/\gamma} \quad (8)$$

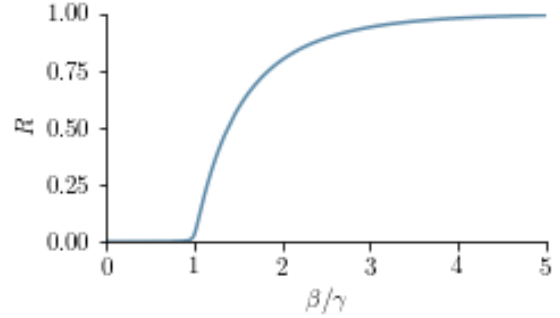


Figura 4: Dimensione dell'epidemia in funzione di β/γ

Possiamo studiare il comportamento di questa funzione graficamente: tracciando le curve $y(R) = 1 - e^{-\beta R/\gamma}$ al variare di β/γ otteniamo l'andamento mostrato in figura 3. La linea tratteggiata corrisponde a $y = R$. In corrispondenza dell'intersezione si ha la soluzione $R = 1 - e^{-\beta R/\gamma}$.

Si può vedere come al variare di β/γ cambia il numero di punti di intersezione della curva con la diagonale: per bassi valori di β/γ le curve si intersecano solo all'origine e la soluzione è $R = 0$; per β/γ sufficientemente grande, invece, le curve si intersecano in due punti, uno a $R = 0$ e uno a $R > 0$.

Se tracciamo la curva corrispondente alle soluzioni così ottenute in funzione di β/γ otteniamo il grafico in figura 4. È immediato vedere come per $\beta/\gamma \leq 1$ l'epidemia non ha modo di scoppiare, infatti, quando $\beta \leq \gamma$ gli individui infetti recuperano più rapidamente di quanto i suscettibili vengano infettati.

Questo comportamento combacia con la definizione di rate riproduttivo R_0 che abbiamo dato nella sezione precedente.

$$R_0 = \frac{\beta}{\gamma} \quad (9)$$

$R_0 = 1$ viene chiamato *soglia di transizione* e demarca il punto in cui l'andamento dell'epidemia passa da crescente a decrescente.

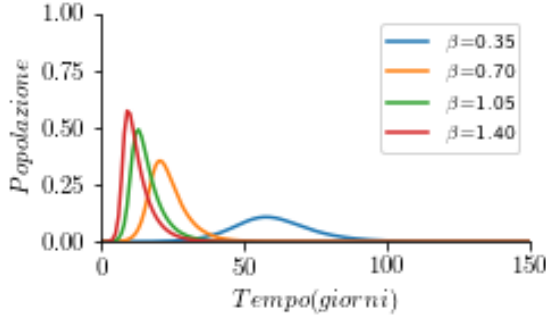


Figura 5: Andamento della curva I al variare di β . $\gamma = 0.35$

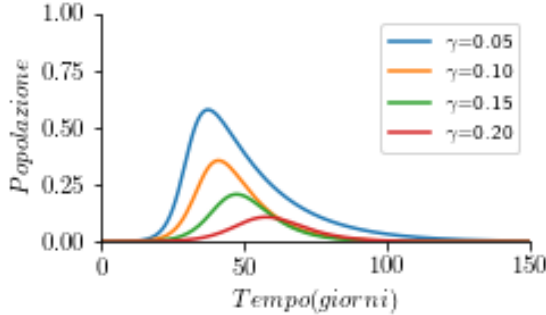


Figura 6: Andamento della curva I al variare di γ . $\beta = 0.35$

I grafici in Figura 5 e 6 mostrano l'andamento della curva di frazione di popolazione Infetta al variare dei fattori β e γ .

All'aumentare di β , il picco della curva è sempre più elevato e si presenta per valori di t inferiori; l'epidemia impiega meno tempo per infettare lo stesso numero di individui e una maggiore porzione della popolazione totale viene infettata nei primi giorni.

Variando γ , invece, possiamo vedere che la posizione nel tempo del picco non varia; decresce la percentuale massima di popolazione infetta e la durata totale dell'epidemia.

2.1 Integrazione del sistema

Per trovare una soluzione delle equazioni differenziali non lineari del sistema non è possibile integrare in modo esplicito, e risulta necessario usare metodi di approssimazione.

Detto Δt un intervallo di tempo e $t_k = k\Delta t$ l'orbita, si può scrivere:

$$\mathbf{x}(t_k) = S_{\Delta t}(\mathbf{x}(t_{k-1})) \quad (10)$$

Una soluzione approssimata si ottiene da un'evoluzione approssimata di ordine m con un resto di ordine $m + 1$.

$$S_{\Delta t} = M_{\Delta t} + o(\Delta t^{m+1}) \quad (11)$$

La mappa $M_{\Delta t}$ genera un'orbita discreta

$$\mathbf{x}_k = M_{\Delta t}(\mathbf{x}_{k-1}) \quad (12)$$

A seguire l'integrazione del modello presentato tramite metodo di Eulero al primo ordine e Runge-Kutta al quarto ordine.

Metodo di Eulero

Il metodo è basato sullo sviluppo di Taylor, dove la mappa M_k si ottiene dallo sviluppo fino all'ordine m trascurando il resto.

Lo schema al primo ordine è dato da

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta t \Phi(\mathbf{x}_k)$$

Con $\Phi(\mathbf{x}_k)$ il sistema differenziale da integrare. Applicato al nostro modello otteniamo le equazioni

$$S_{k+1} = S_k - \Delta t \beta S_k I_k \quad (13a)$$

$$I_{k+1} = I_k + \Delta t \beta S_k I_k - \Delta t \gamma I_k \quad (13b)$$

$$R_{k+1} = R_k + \Delta t \gamma I_k \quad (13c)$$

Dalle equazioni 13 è possibile ricavare grafici analoghi a quelli in Figura 7.

L'errore locale commesso dal metodo di eulero si può stimare sottraendo la soluzione numerica a un dato step, x_k , alla soluzione esatta al tempo successivo $t_k + 1 = t_k + \Delta t$ ottenuta dallo sviluppo di Taylor.

$$\begin{aligned} \|E\| &= \mathbf{x}(t_k + \Delta t) - \mathbf{x}_k = \\ &= \frac{1}{2} \Delta t^2 \mathbf{x}''_k + o(\Delta t^3) \end{aligned}$$

L'errore locale risulta quindi proporzionale a Δt^2 .

L'errore globale a un dato tempo t , invece, è definito come l'effetto accumulato degli errori locali commessi negli step necessari per arrivare al dato tempo. In particolare l'errore globale è

$$e_k \leq C \Delta t (e^{Lt_k} - 1)$$

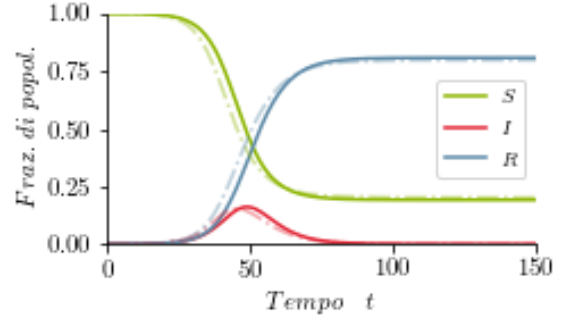
dove L è la costante di Lipschitz della funzione. Quando $L \rightarrow 0$ gli errori semplicemente si sommano e $e_k \leq C \Delta t^2$.

Essendo $L > 0$ si ha un effetto di amplificazione e l'errore cresce esponenzialmente per alti valori di t .

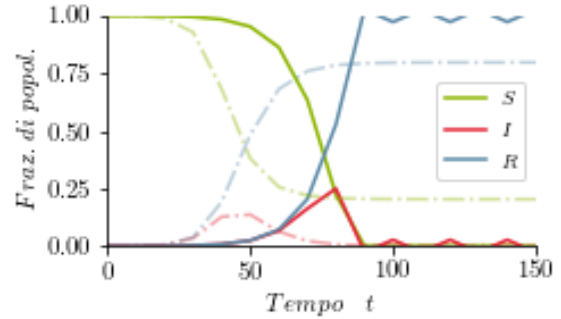
Questo risultato è evidente per $\Delta t > 1$ come nel caso di Figura 7b, dove l'errore è molto alto e la soluzione approssimata risulta numericamente instabile.

Metodo di Runge-Kutta

Per ridurre l'errore sarebbe conveniente effettuare il calcolo con un metodo di ordine più



(a) $\Delta t = 1$



(b) $\Delta t = 10$

Figura 7: Integrazione con il metodo di Eulero. I parametri sono $\beta = 0.4$ e $\gamma = 0.2$. Le linee tratteggiate corrispondono alla soluzione "esatta"¹.

elevato, ma il metodo di Eulero richiede il calcolo di derivate che risulterebbe oneroso per gradi elevati.

Si può quindi utilizzare lo schema noto come metodo Runge-Kutta. Esso prevede il calcolo

¹La soluzione è ottenuta tramite l'algoritmo LSODA della libreria Odepack di FORTRAN. L'algoritmo alterna automaticamente tra metodi di Adams-Bashfort e Formule di differenziazione Backward al fine di ottimizzare il risultato.[5]

di un sistema algebrico così definito:

$$\begin{cases} K_1 = \Phi(x_k, t_k) \\ K_2 = \Phi(x_k + \frac{\Delta t}{2} K_1, t_k + \frac{\Delta t}{2}) \\ K_3 = \Phi(x_k + \frac{\Delta t}{2} K_2, t_k + \frac{\Delta t}{2}) \\ K_4 = \Phi(x_k + \Delta t K_3, t_k + \Delta t) \\ x_{k+1} = x_k + \frac{\Delta t}{6} (K_1 + 2K_2 + 2K_3 + K_4) \end{cases} \quad (14)$$

Nel nostro caso Φ corrisponde al sistema differenziale definito in 3 discretizzato seguendo le definizioni di equazione 10, 11 e 12.

Il metodo è uno schema di quart'ordine. L'errore viene studiato in maniera analoga al metodo di Eulero, quindi l'errore globale è dell'ordine di $o(\Delta t^5)$ e l'errore globale $o(\Delta t^4)$.

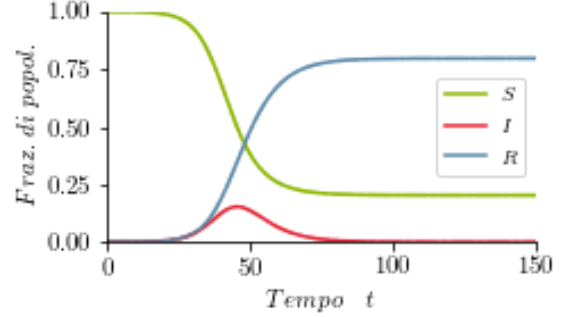
La Figura 8 mostra le curve risultanti dall'integrazione con il metodo di Runge-Kutta per differenti valori di Δt . Figura 9 mostra un dettaglio di Figura 8b; si vede come con il metodo di Runge-Kutta, anche per Δt alto, l'errore è significativamente basso.

Le Appendici A e B presentano gli algoritmi in codice Python sviluppati per costruire gli integratori descritti in questa sezione.

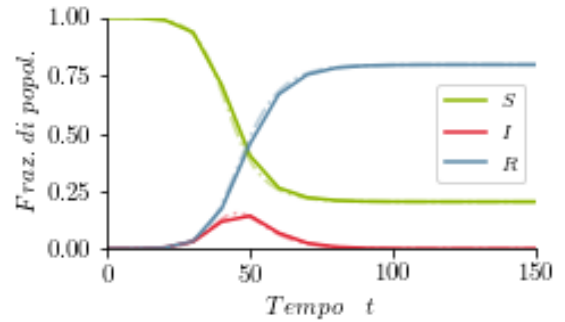
3 Cellular Automata

Gli *automi cellulari* (CA) sono modelli di sistemi dinamici, con spazio e tempo discreti ed interazioni esclusivamente locali.

Un automa cellulare consiste in un reticolo d-dimensionale uniforme. Ad ogni cella del reticolo corrisponde un valore, lo stato della cella. Ogni cella è limitata a interazioni locali con i soli vicini, ed è quindi incapace di “comunicare” globalmente con le altre celle del reticolo. Lo stato di ogni cella è aggiornato simultaneamente ad intervalli discreti di tempo, in accordo con una regola di transizione che tiene conto dello stato della cella e delle sue vicine all'intervallo di tempo precedente. Solitamente ad ogni cella è applicata la stessa regola di transizione.



(a) $\Delta t = 1$



(b) $\Delta t = 10$

Figura 8: Integrazione con il metodo di Runge-Kutta di ordine 4. I parametri sono $\beta = 0.4$ e $\gamma = 0.2$. Le linee tratteggiate corrispondono alla soluzione “esatta”.

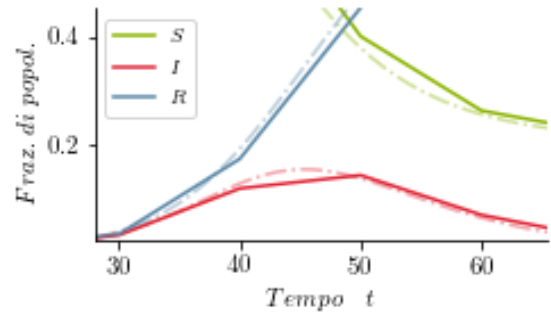


Figura 9: Dettaglio di Figura 8b

Formalmente, un automa cellulare è definito dalla quadrupla (C, Q, N, f) , dove:

- C è il reticolo cellulare a d dimensioni con $C = \{(i_1, \dots, i_d); 1 \leq i_j \leq w_j\}$ dove w_j è la larghezza del reticolo nella j -esima dimensione;
- Q è lo spazio degli stati, l'insieme finito e discreto dei possibili stati q delle celle del reticolo;
- N è l'indice di vicinato, un insieme finito di indici $N = (\alpha_1, \dots, \alpha_d) \subset \mathbb{Z}^d$ che definisce per ogni cella (i_1, \dots, i_d) i suoi n vicini:

$$N_{(i_1, \dots, i_d)} = \{(i_1 + \alpha_{11}, \dots, i_d + \alpha_{d1}), \dots, (i_1 + \alpha_{1n}, \dots, i_d + \alpha_{dn})\}$$

- f è la regola di transizione, una funzione $f : Q \rightarrow Q$ tale che, detto q_i^t lo stato della cella nel punto $i = (i_1, \dots, i_d)$ dello spazio C al tempo t , e indicati con $(\alpha_1, \dots, \alpha_n)$ gli elementi di N_i risulta:

$$q_i^t = f(q_{i+\alpha_1}^{t-1}, \dots, q_{i+\alpha_n}^{t-1})$$

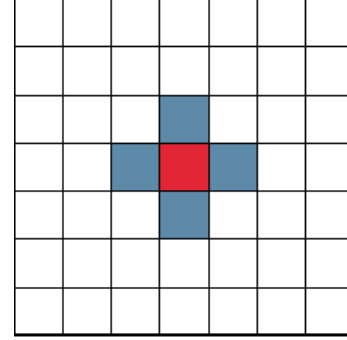
Nell'esempio classico a 2 dimensioni, il CA è composto da un reticolo:

$$C_{M \times N} = \{(x, y); 1 \leq x \leq M, 1 \leq y \leq N\}$$

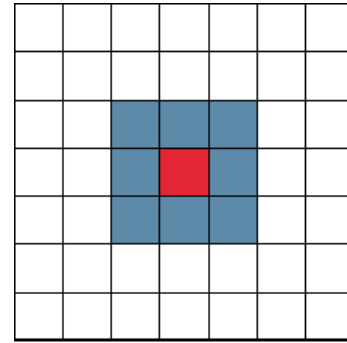
Per i reticoli bidimensionali sono definiti, solitamente, due diversi vicini:

- Il vicinato di *von Neumann*, definito dagli indici
 $N = \{(0, 0), (0, 1), (1, 0), (0, -1), (-1, 0)\}$
- Il vicinato di *Moore*, definito dagli indici
 $N = \{(0, 0), (0, 1), (1, 1), (1, 0), (1, -1), (0, -1), (-1, -1), (-1, 0), (-1, 1)\}$

In Figura 10 sono rappresentati dei CA bidimensionali con reticolo 7×7 ; sono evidenziate in rosso la cella e in blu le sue vicine secondo von Neumann (10a) e Moore 10b.



(a)



(b)

Figura 10: Cellular automata bidimensionale. In rosso è evidenziata la cella e in blu le vicine secondo *von Neumann* (a) e secondo *Moore* (b).

3.1 Modello Proposto

Per il modello SIR in analisi useremo un CA bidimensionale, con griglia $r \times c$ con r il numero di righe e c il numero di colonne.

$$C = \{(i, j); 0 \leq i \leq r-1, 0 \leq j \leq c-1\}$$

Il reticolo C rappresenta il territorio su cui si diffonde la malattia. Ogni cella corrisponde a un'area di popolazione costante divisa nelle tre classi descritte in Sezione 1.

Tenendo conto della discretizzazione dello spazio e del tempo definiamo

$$0 \leq S_{i,j}^t \leq 1, 0 \leq I_{i,j}^t \leq 1, 0 \leq R_{i,j}^t \leq 1$$

con $S_{i,j}^t + I_{i,j}^t + R_{i,j}^t = 1$

le frazioni di popolazione di ogni cella i, j al tempo t .

Per ottenere lo spazio Q di stati discreti definiamo

$$\begin{aligned}\tilde{S}_{i,j}^t &= \frac{\lfloor K \cdot S_{i,j}^t \rfloor}{K}, \\ \tilde{I}_{i,j}^t &= \frac{\lfloor K \cdot I_{i,j}^t \rfloor}{K}, \\ \tilde{R}_{i,j}^t &= \frac{\lfloor K \cdot R_{i,j}^t \rfloor}{K}\end{aligned}$$

Dove $\lfloor \cdot \rfloor$ è l'operazione Parte Intera. Con $K = 100$ l'insieme dei possibili valori di stato delle celle è $E = \{0, 0.01, 0.02, \dots, 0.99, 1\}$ e

$$\begin{aligned}q_{i,j}^t &= (\tilde{S}_{i,j}^t, \tilde{I}_{i,j}^t, \tilde{R}_{i,j}^t) \\ \text{con } q_{i,j}^t &\in Q = E^3\end{aligned}$$

La regola di transizione per ogni cella è data dal sistema:

$$S_{i,j}^{t+1} = S_{i,j}^t - \sum_{(u,v) \in N_{i,j}} \beta \nu_{i,j}^{u,v} S_{i,j}^t I_{u,v}^t \quad (15a)$$

$$I_{i,j}^{t+1} = (1 - \gamma) I_{i,j}^t + \sum_{(u,v) \in N_{i,j}} \beta \nu_{i,j}^{u,v} S_{i,j}^t I_{u,v}^t \quad (15b)$$

$$R_{i,j}^{t+1} = R_{i,j}^t + \gamma I_{i,j}^t \quad (15c)$$

con β l'infettività della malattia, γ il tasso di recupero e ν il rate di contatto, cioè la porzione di contatti effettivi tra individui della cella (i, j) e del suo vicino (u, v) .

Il caso $(u, v) = (i, j)$ rappresenta il contatto tra individui della stessa cella.

Il modello così rappresentato ha un comportamento simile a quello rappresentato dal diagramma in Figura 1, con la differenza che la popolazione si considera a contatto con i soli vicini N , inclusa la cella stessa.

Per determinare il possibile scoppio dell'epidemia consideriamo le celle $O = (i, j)$ e una sua vicina $A = (u, v) \in N_{(i,j)} - \{(i, j)\}$.

β	ν	ρ
0.3	0.3	0.11
	0.6	0.06
	1	0.03
0.6	0.3	0.06
	0.6	0.03
	1	0.02
1	0.3	0.03
	0.6	0.02
	1	0.01

Tabella 1: Porzione minima di popolazione infetta nella cella di origine al tempo $t = 0$ perché l'infezione si diffonda alle celle adiacenti, in funzione di β e ν

Supponiamo che al tempo $t = 0$ vi sia la sola cella O con popolazione infetta nel reticolo e le restanti celle abbiano popolazione di soli suscettibili.

Dall'equazione 15b la popolazione infetta della cella A al tempo $t = 1$ risulta:

$$I_A^1 = \beta \cdot \nu_A^O \cdot I_O^0$$

Considerando stati discreti avremo una popolazione infetta nella cella A al tempo $t + 1$ se $\tilde{I}_A^1 \in E - \{0\}$, cioè $\tilde{I}_A^1 \geq 0.01$. Di conseguenza avremo:

$$\tilde{I}_O^0 \geq \frac{1}{100 \cdot \beta \cdot \nu_A^O} = \rho \quad (16)$$

che corrisponde alla porzione di individui infetti nella cella d'origine dell'infezione necessaria perché l'epidemia si diffonda alle celle vicine.

Il valore di ρ dipende da l'infettività β della malattia e dal rate di contatto ν . (Tabella 1)

Per valori di $\tilde{I}_O^0 \leq \rho$ l'epidemia non si diffonde alle celle vicine e il numero di infetti tende a zero se $I_O^{t+1} < I_O^t \forall t$. Quindi, considerando equazione 15b con $(u, v) = (i, j) = O$:

$$I_O^{t+1} = (1 - \gamma) I_O^t + \beta \nu_O^O S_O^t I_O^t < I_O^t$$

che implica:

$$\frac{\beta \cdot \nu_O^O \cdot S_O^t}{\gamma} < 1 \quad (17)$$

che corrisponde al rate riproduttivo R_0 come definito nella sezione precedente, quando $\nu_O^O \cdot S_O^t \sim 1$.

Quando l'equazione 17 non è valida, la popolazione infetta della cella aumenta, diffondendosi alle celle vicine una volta superato il valore ρ .

3.2 Andamento spaziale

La simulazione del modello è stata svolta con i seguenti parametri: $r = 50$, $c = 50$, $\beta = 0.3$, $\gamma = 0.05$; $\nu_{i,j}^{u,v} = 0.5$ e $\nu_{i,j}^{i,j} = 1$. La malattia emerge al tempo $t = 0$ nella cella $(25, 25)$ con $\tilde{S}_{i,j}^t = 0.5$, $\tilde{I}_{i,j}^t = 0.5$, $\tilde{R}_{i,j}^t = 0$.

Entrambi i tipi di vicinato descritti sono usati.

Nel caso del vicinato di Moore (Figura 11) l'epidemia si diffonde più rapidamente, copre una superficie maggiore e raggiunge i limiti del reticolo prima che nel caso del vicinato alla VonNeumann (Figura 12)

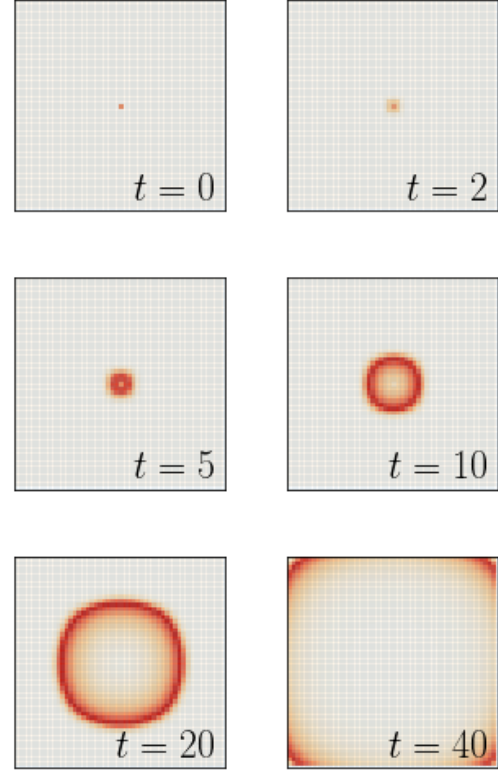
I risultati ottenuti mostrano come la struttura del vicinato influenzi significativamente l'evoluzione dell'epidemia, determinando il picco massimo di infezione e la sua durata.

Si può notare inoltre un problema di saturazione nella salita della curva di popolazione Infetta, dovuto alla località delle interazioni. Ogni cella infatti può interagire con le sole vicine, e quando la porzione di popolazione infetta nelle vicine raggiunge 1 non vi è più contatto efficace dalla cella di origine.

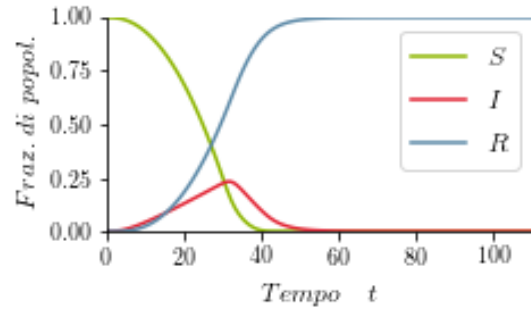
Per ridurre l'effetto possiamo estendere l'insieme di vicini tramite il concetto di distanza di Čebyšëv.

$$N_{(i,j)} = \{(u,v); \max(|u-i|, |v-j|) \leq r\}$$

rappresentato in Figura 13.



(a)



(b)

Figura 11: Reticolo C a diversi istanti di tempo con indice di vicinato di Moore (a) e le rispettive curve di popolazione risultanti (b).

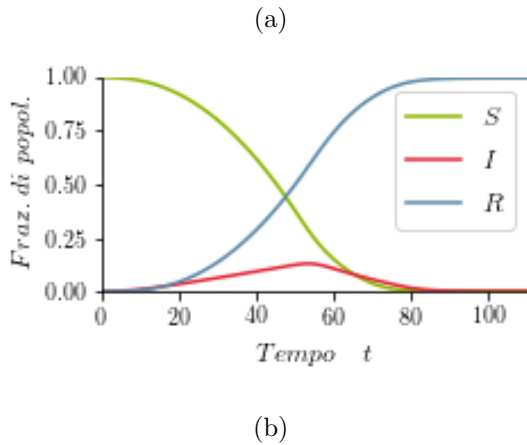
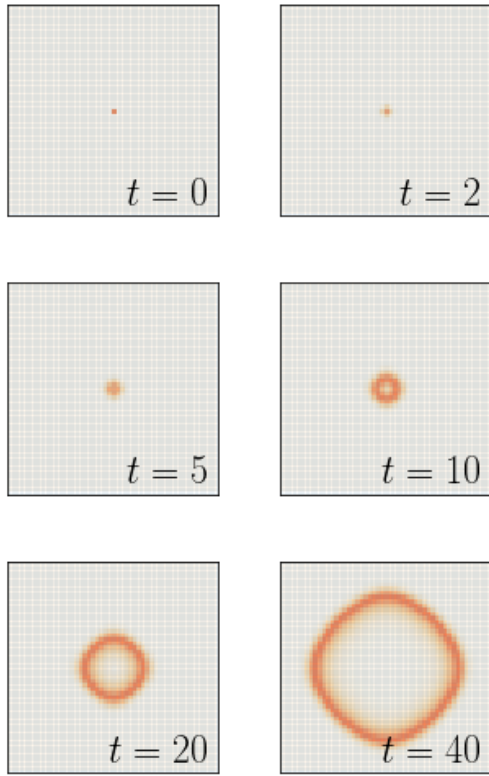


Figura 12: Reticolo C a diversi istanti di tempo con indice di vicinato di VonNeumann (a) e le rispettive curve di popolazione risultanti (b).

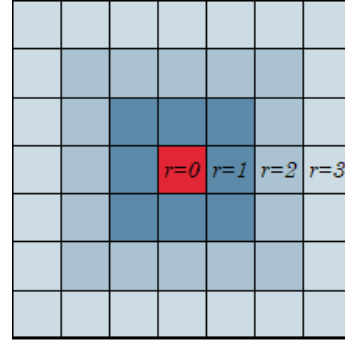


Figura 13: Vicinato secondo la distanza di Čebyšëv con raggio r

Per la simulazione utilizziamo i parametri impostati precedentemente, variando esclusivamente il rate di contatto in funzione della distanza r del vicino (u, v) :

$$\nu_{i,j}^{u,v} = \frac{1}{r+1}$$

Il risultato è visibile in Figura 14.

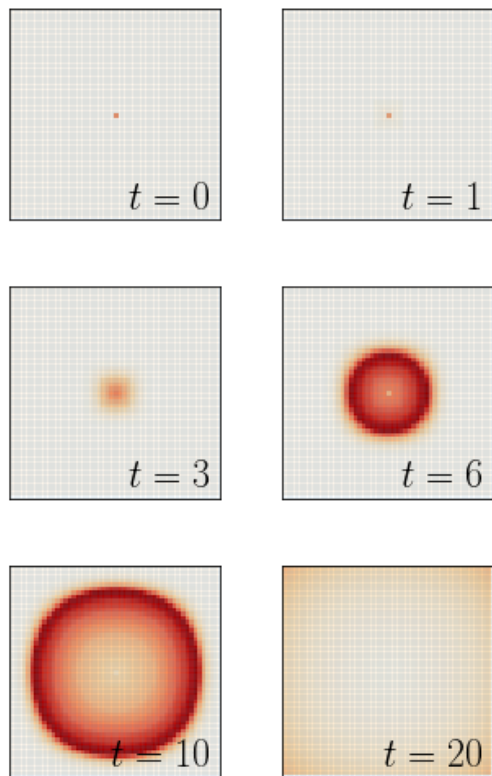
In Appendice C è trascritta parte del codice Python della simulazione presentata in questa sezione.

4 Conclusione

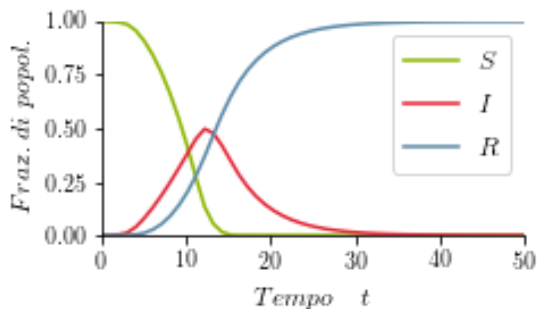
In questo paper si è analizzato un modello compartimentale di diffusione dell'epidemie, il modello SIR.

Il modello prevede la diffusione del morbo tramite contatto diretto tra individui infetti e suscettibili, con una durata sufficientemente breve da trascurare immigrazioni, emigrazioni, nascite e morti, garantendo quindi una popolazione totale costante (Sezione 1).

La simulazione del modello è stata svolta con due differenti metodi: il più classico metodo differenziale, descritto in sezione 2, e una simulazione tramite un automa cellulare bidimensionale, che viene descritto in sezione 3.



(a)



(b)

Figura 14: Reticolo C a diversi istanti di tempo con indice di vicinato di Čebyšëv (a) e le rispettive curve di popolazione risultanti (b).

I due modelli, pur presentando andamenti simili hanno caratteristiche differenti che li rendono più o meno efficaci in diversi ambiti.

Il metodo differenziale prevede che le classi di popolazione siano omogeneamente distribuite e che sia possibile il contatto tra un individuo ed ogni altro individuo della popolazione.

Questo metodo, quindi, presenta il vantaggio di avere una dinamica più semplice da analizzare, ma non tiene conto della spazialità dell'infezione.

Il secondo metodo consiste, invece, nell'analisi del modello tramite lo studio dell'andamento di un automa cellulare bidimensionale, con stati tridimensionali e opportune regole di transizione, analoghe alle equazioni differenziali che descrivono il modello differenziale.

Il metodo tiene ovviamente conto della spazialità del problema ed è evidente dai grafici di confronto fra i differenti modelli di vicinato (Figure 11, 12, 14) come le differenti regole di contatto influenzino significativamente l'andamento dell'epidemia e il picco massimo di infetti.

Per il modello è stato studiato in entrambi i casi il valore soglia dello scoppio di epidemia, calcolando per il modello CA, sia che valori di popolazione iniziale sono necessari perché l'infezione si diffonda alle celle vicine, sia che rate di infettività e contatto sono necessari perché il numero degli infetti aumenti prima di decrescere.

È possibile ricondursi a un comportamento paragonabile al caso differenziale considerando come vicinato di una cella l'intero reticolo, scegliendo celle random come contatti efficaci, quindi in grado di diffondere la malattia.

Un comportamento simile è paragonabile a una teoria di campo medio, e il risultato prodotto combacia effettivamente con il risultato ottenuto dal modello differenziale conside-

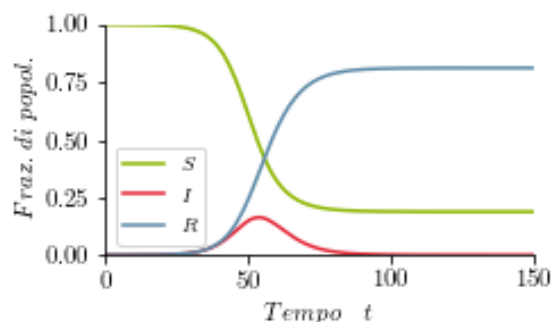


Figura 15: Curve di popolazione S, I e R ottenute considerando contatti casuali all'interno dell'intera popolazione.

rando una popolazione infetta a contatto con l'intera popolazione suscettibile.

Figura 15 mostra le curve di popolazione ottenute con questo metodo.

Riferimenti bibliografici

- [1] Amine Bouaine e Mostafa Rachik. "Modeling the impact of immigration and climatic conditions on the epidemic spreading based on cellular automata approach". In: *Ecological Informatics* 46 (2018), pp. 36–44. ISSN: 1574-9541.
- [2] Martin Gardner. "The fantastic combinations of John Conway's new solitaire game" life" by Martin Gardner". In: *Scientific American*. Mathematical games 223 (1970), pp. 120–123.
- [3] Herbert W Hethcote. "Qualitative analyses of communicable disease models". In: *Mathematical Biosciences* 28.3-4 (1976), pp. 335–356.
- [4] Herbert W Hethcote e James A Yorke. "Gonorrhea Transmission Dynamics and Control". In: *Lecture notes in biomathematics* 56 (1984).
- [5] Alan C Hindmarsh. "ODEPACK, a systematized collection of ODE solvers". In: *Scientific computing* (1983).
- [6] William O Kermack e Anderson G McKendrick. "Contributions to the mathematical theory of epidemics. II.—The problem of endemicity". In: *Proc. R. Soc. Lond. A* 138.834 (1932), pp. 55–83.
- [7] Stanislaw Ulam. "John von Neumann 1903-1957". In: *Bulletin of the American mathematical society* 64.3 (1958), pp. 1–49.
- [8] S Hoya White, A Martin Del Rey e G Rodriguez Sanchez. "Modeling epidemics using cellular automata". In: *Applied Mathematics and Computation* 186.1 (2007), pp. 193–202.
- [9] Stephen Wolfram. *A new kind of science*. Vol. 5. Wolfram media Champaign, 2002.

Appendici

Le integrazioni e le simulazioni descritte nell'elaborato sono state sviluppate in Python 3, con il supporto della libreria numerica Numpy e le librerie Pylab, Matplotlib e Seaborn per i grafici.

Le sezioni seguenti presentano estratti dei codici Python utilizzati nel paper. In particolare sono state trascritti i soli algoritmi, escludendo le routine sviluppate per la realizzazione dei grafici mostrati nell'elaborato.

A Eulero

```
1 def derivate(x, beta, gamma):
2     S, I, R = x
3     dSdt = -beta * S * I
4     dIdt = beta * S * I - gamma * I
5     dRdt = gamma * I
6     return dSdt, dIdt, dRdt
7
8 E[0] = S0, I0, R0
9
10 for i in range(1, len(t)):
11     E[i] = E[i - 1] + derivate(E[i], beta, gamma) * dt
```

L'integratore di Eulero contiene una singola funzione (`derivate`), corrispondente al sistema in equazione 3 di sezione 2.

L'array `E[0]`, inizializzato in riga 8, viene riempito nel ciclo di riga 10 seguendo il sistema di equazioni 13 in sezione 2.

Non trascritte nel codice sopra, sono le inizializzazioni dei valori delle classi al tempo 0, i parametri `beta` e `gamma` e l'intervallo di tempo `dt`.

B Runge-Kutta

```
1 def derivate(y, t, beta, gamma):
2     S, I, R = y
3     dSdt = -beta * S * I
4     dIdt = beta * S * I - gamma * I
5     dRdt = gamma * I
6     return np.array([dSdt, dIdt, dRdt])
7
8 def RungeKutta(x_k, t_k, beta, gamma, dt):
9     S_k, I_k, R_k = x_k
10    K1 = derivate(x_k, t_k, beta, gamma)
11    K2 = derivate(x_k+dt/2*K1, t_k+dt/2, beta, gamma)
```

```

12     K3 = derivate(x_k+dt/2*K2, t_k+dt/2, beta, gamma)
13     K4 = derivate(x_k+dt*K3, t_k+dt, beta, gamma)
14     return x_k+dt/6*(K1+2*K2+2*K3+K4)
15
16 RK[0] = S0, I0, R0
17
18 for i in range(1, len(t)):
19     RK[i] = RungeKutta(RK[i-1], i-1, beta, gamma, dt)

```

L'integratore alla Runge-Kutta presenta due funzioni principali.

La funzione `derivate` è del tutto analoga a quella presente nell'algoritmo del metodo di Eulero in Appendice A, con l'unica differenza nel metodo di riporto in riga 6, dove il risultato è passato come `Numpy.array`, tipo che consente il prodotto di uno scalare per un vettore.

La seconda funzione (`RungeKutta`), segue il sistema 14 di sezione 2.

In riga 16 è inizializzato, ai valori `S0`, `I0`, `R0` al tempo 0 delle frazioni di popolazione nelle rispettive classi, il vettore contenente le tre curve del sistema.

Il ciclo nelle ultime due righe, richiama la funzione `RungeKutta` che riempie la lista delle soluzioni.

Nuovamente sono state omesse nella trascrizione le inizializzazioni delle variabili in gioco.

C Automa Cellulare

```

1 def hoodSum(i, j, I):
2     result = 0
3     for Neighbor in Hood:
4         a = i + Neighbor[0]
5         b = j + Neighbor[1]
6         if a >= CAside or b >= CAside or a < 0 or b < 0:
7             continue
8         sum = sum + nu * I[a, b]
9     return sum
10
11 def updateMatrices(x):
12     S, I, R = x
13     for i in range(CAside):
14         for j in range(CAside):
15             nR[i, j] = min(R[i, j] + gamma * I[i, j], 1)
16             nI[i, j] = max(min((1. - gamma) * I[i, j] + beta * S[i, j] *
17                               ↪ hoodSum(i, j, I), 1), 0)
18             nS[i, j] = 1. - nR[i, j] - nI[i, j]
19     return nS.copy(), nI.copy(), nR.copy()
20 Hood = [[1, 0], [0, 1], [0, -1], [-1, 0], [0, 0]]

```

```
21
22 CA[0] = S0, I0, R0
23
24 for t in range(time):
25     CA.append(updateMatrices(CA[t]))
```

L'algoritmo, che calcola lo sviluppo dell'automa cellulare presentato in sezione 3, è composto dalla funzione `updateMatrices`, che calcola lo stato a 3 dimensioni di ogni cella del reticolo bidimensionale, e dalla funzione `hoodSum` che svolge la sommatoria presente nelle equazioni 15b.

Tramite il doppio ciclo (riga 13-18) vengono costruiti tre nuove matrici, corrispondenti agli stati (`nS`, `nI`, `nR`) delle celle i, j al tempo $t+1$.

Le funzioni `min` e `max` in riga 15 e 16, assicurano che i valori delle frazioni di popolazione siano compresi tra 0 e 1.

La lista `Hood[...]` in riga 20 definisce gli indici di vicinato utilizzati; nel caso trascritto corrispondono al vicinato di VonNeumann.

La lista contenente il reticolo ai diversi istanti di tempo è inizializzata in riga 22 e il ciclo nelle ultime righe *appende* in coda alla lista il nuovo stato del CA per ogni t .