



993SM - Laboratory of Computational Physics lecture 14 - part 1 June 10, 2020

Maria Peressi

Università degli Studi di Trieste - Dipartimento di Fisica
Sede di Miramare (Strada Costiera 11, Trieste)

e-mail: peressi@ts.infn.it

tel.: +39 040 2240242

Modelling other random processes

- genetic algorithms
- Stochastic models (prey-predator and epidemiologic) and deterministic version
- Lotka-Volterra, Wa-Tor, Reed-Frost...

M. Peressi - UniTS - Laurea Magistrale in Physics
Laboratory of Computational Physics - Unit XIV

Algoritmi genetici

un problema di ottimizzazione

idea: applicare i principi dell'evoluzionismo naturale a sistemi artificiali, codificando in modo numerico configurazioni di input, processi evolutivi, soluzioni, etc etc

Un po' di terminologia:

- popolazione
- genoma
- fitness
- processi di:
crossover, mutazione,

Popolazione
(1 riga = 1 individuo)

genoma

0	1	0	1	0	1	0	1
1	0	1	1	0	0	0	0
1	0	1	1	0	1	1	0

1. Generazione casuale di una popolazione iniziale costituita da un certo numero di “individui”, ognuno con un certo “genoma”;
- 2.
- 3.

Popolazione
(1 riga = 1 individuo)

genoma

fitness

0	1	0	1	0	1	0	1	4
1	0	1	1	0	0	0	0	3
1	0	1	1	0	1	1	0	5

1. Generazione casuale di una popolazione iniziale costituita da un certo numero di “individui”, ognuno con un certo “genoma”;
2. ciascun individuo ha una propria **fitness**, che è un indice della qualità della “soluzione” che egli rappresenta; es: $0+1+0+1+0+1+0+1=4$
- 3.

genoma

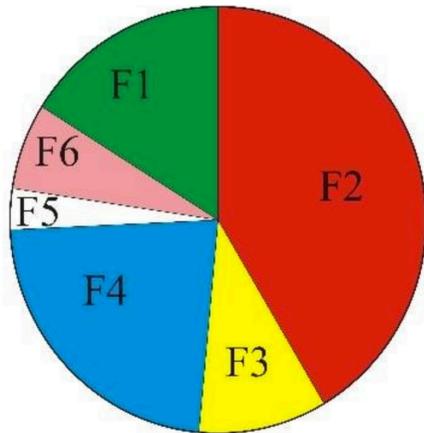
genitore	0	1	0	1	0	1	0	1
	1	0	1	1	0	0	0	0
genitore	1	0	1	1	0	1	1	0

fitness

4

3

5



1. Generazione casuale di una popolazione iniziale costituita da un certo numero di “individui”, ognuno con un certo “genoma”;
2. ciascun individuo ha una propria **fitness**, che è un indice della qualità della “soluzione” che egli rappresenta; es: $0+1+0+1+0+1+0+1=4$
3. Inizio di ciclo evolutivo ha inizio: la **selezione** simula la selezione naturale darwiniana. Ogni individuo (potenziale genitore) ha una probabilità di essere selezionato proporzionale alla fitness (\Rightarrow *roulette wheel*). I selezionati sono considerati “genitori” e danno luogo a “nuovi individui”.

	genoma								fitness
genitore	0	1	0	1	0	1	0	1	4
genitore	1	0	1	0	0	0	0	0	3
genitore	1	0	1	1	0	1	1	0	5

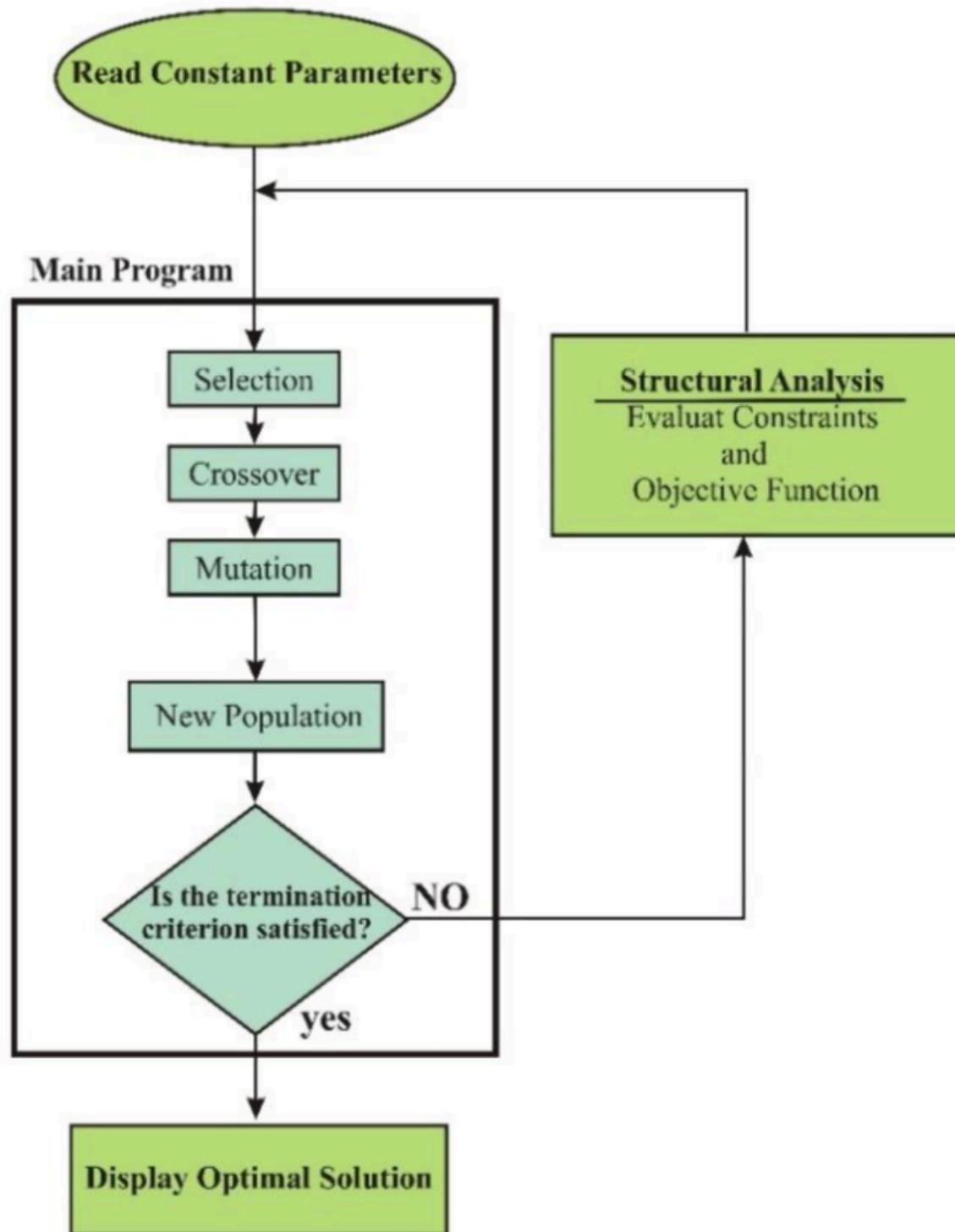
Dopo il crossover:

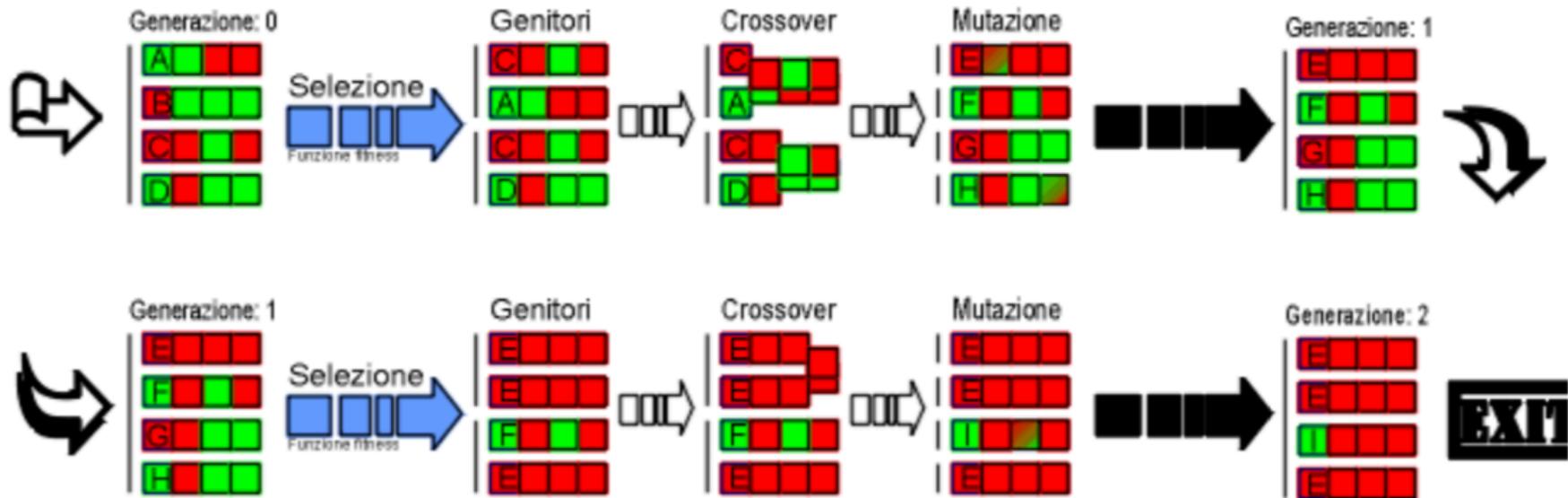
0	1	1	1	0	1	0	1
1	0	0	1	0	1	0	1

Dopo la mutazione:

0	1	1	1	0	1	0	1
1	0	0	1	1	1	0	1

- la ricombinazione (*crossover*) agisce sulla popolazione intermedia dei genitori accoppiandoli due a due e scambiandone porzioni di DNA;
- Una *mutazione* può cambiare singoli elementi costitutivi del filamento di DNA e li muta in nuovi;
- l'algoritmo ricomincia e può procedere anche fino all'infinito: si decide un certo criterio di stop (soddisfatta una certa richiesta, ad es. Una certa fitness media)

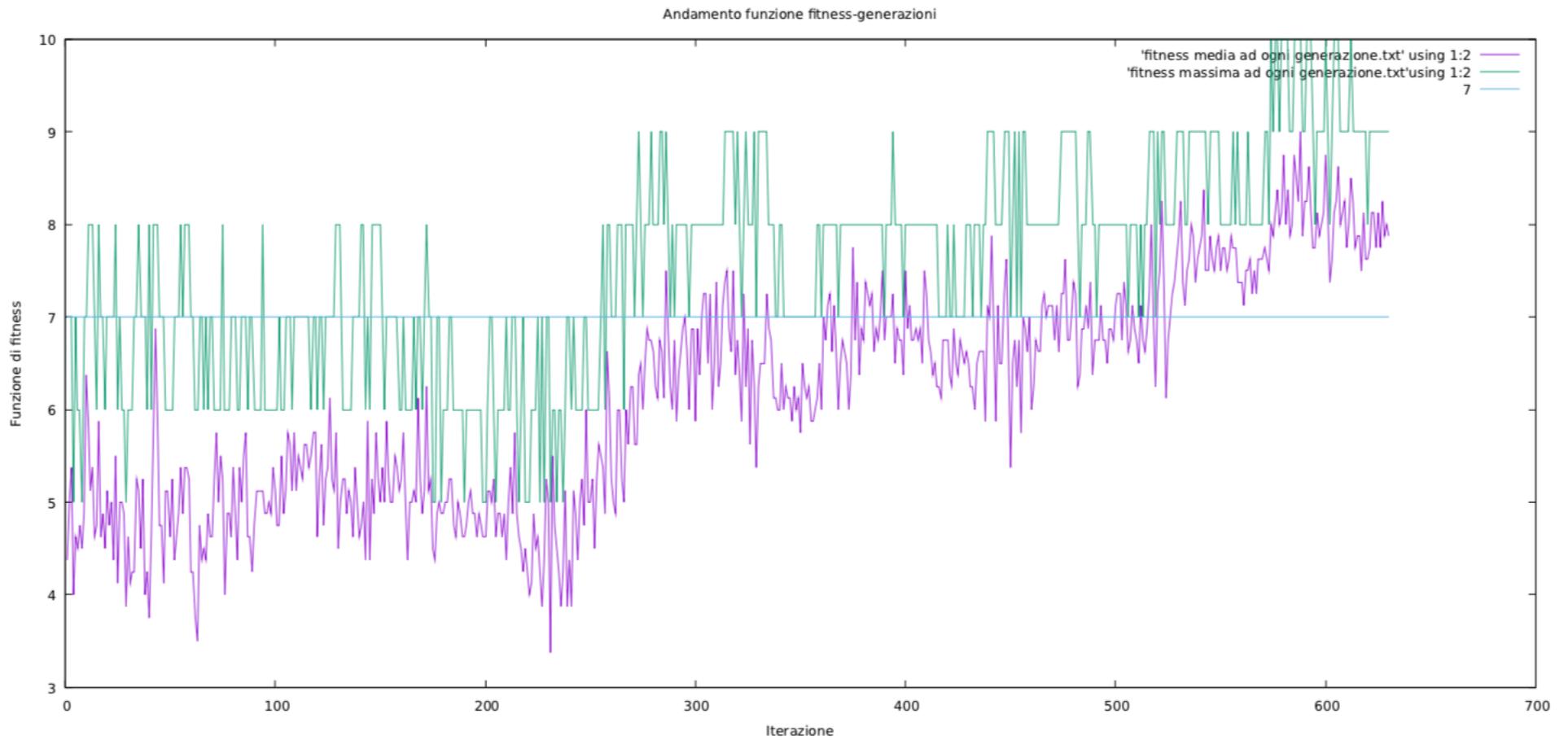




Non esistono dei risultati teorici certi riguardo al numero ottimale della popolazione per un determinato problema, così come non è chiaro quale sia il metodo migliore per scegliere il valore delle probabilità di crossover e mutazione. In genere si procede per via empirica a seconda delle caratteristiche del particolare problema.

PROVARE TEST CON:

- popolazione: da 100 a 10000 (dipende molto dal problema);
- - P_{cross} : circa 0.8;
- - P_{mut} : attorno a 0.01.



#popolazione iniziale = 10, #genitori = 8, Pcross=0.6. Pmut=0.001

Modello deterministico prede-predatori Lotka-Volterra

sistema di equazioni differenziali non lineari del prim'ordine, accoppiate tra loro.

applicazione più famosa: spiegazione della dinamica evolutiva di un sistema prede-predatori

1925 da Alfred J. Lotka e nel 1926 da Vito Volterra:

comprendere l'andamento di alcune statistiche di pesca dei porti del Nord Adriatico relative al periodo 1905-1923.

osservato un anomalo aumento di pesci “grandi”, cioè predatori, durante gli anni della guerra e in quelli immediatamente successivi; siccome una spiegazione imputabile al solo aumento dell'attività umana, quindi da fattori esterni all'ecosistema era insoddisfacente, si tentò una strada alternativa con una modellizzazione.

$$\begin{cases} \frac{dx(t)}{dt} = Ax(t) - Bx(t)y(t) \\ \frac{dy(t)}{dt} = -Cy(t) + Dx(t)y(t) \end{cases}$$

- $A, B, C, D > 0$;
- $x(t)$ è la popolazione della specie predata. In assenza di predatori, essa cresce secondo un modello malthusiano, cioè aumenta in modo incontrollato (modello poco realistico, ipotesi di risorse e capacità riproduttive infinite);
- $y(t)$ è la popolazione della specie predatrice. In assenza di prede essa diminuisce fino all'estinzione (questo è il motivo per cui C è preceduto da un meno);
- $x(t)y(t)$ è la probabilità di incontro tra le due specie: i coefficienti B e D tuttavia non sono uguali tra loro e rappresentano rispettivamente la probabilità di fuga e di cattura;
- essendo un modello di popolazione, $x(t) > 0$ ed $y(t) > 0$.

Il sistema è deterministico, è quindi possibile utilizzare gli strumenti dell'analisi per identificarne alcune caratteristiche peculiari.

Risolvendo le due equazioni omogenee:
$$\begin{cases} \frac{dx(t)}{dt} = 0 \\ \frac{dy(t)}{dt} = 0 \end{cases}$$

è possibile identificare due punti particolari, detti punti fissi.

il primo punto corrispondente all'estinzione di entrambe le specie

secondo ?

Si puo' fare un grafico vettoriale

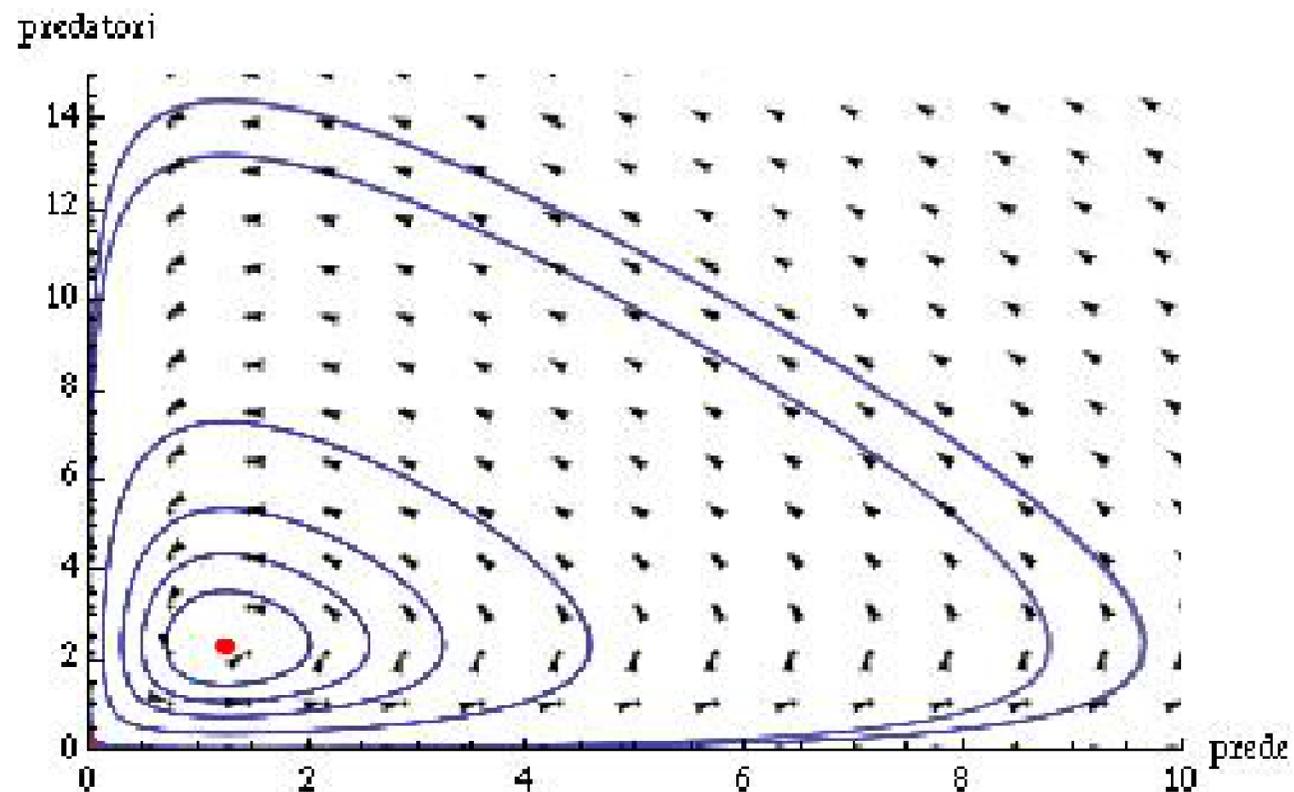


Illustrazione 2: Campo vettoriale nello spazio prede-predatori; sono evidenziati i punti fissi in rosso e alcune orbite del sistema, ottenute tramite integrazioni numerica delle equazioni

Le popolazioni di prede e predatori evolvono dunque in modo ciclico

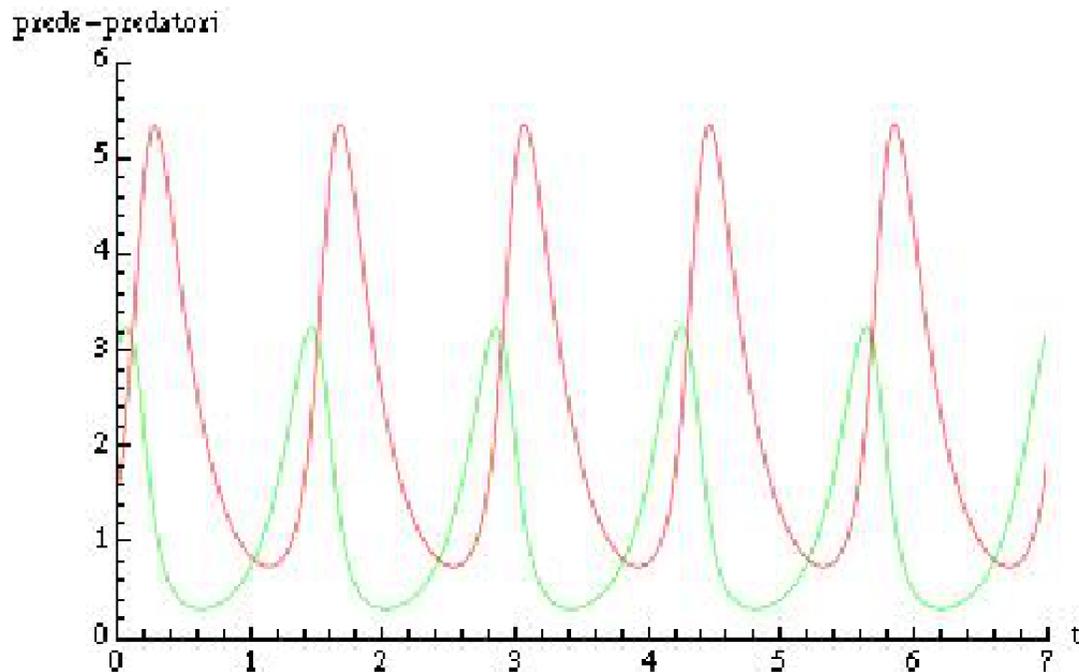


Illustrazione 3: Andamento delle popolazioni di prede (verde) - predatori (rosso) vs. tempo, nel caso di grandi oscillazioni (orbita più esterna del grafico precedente).

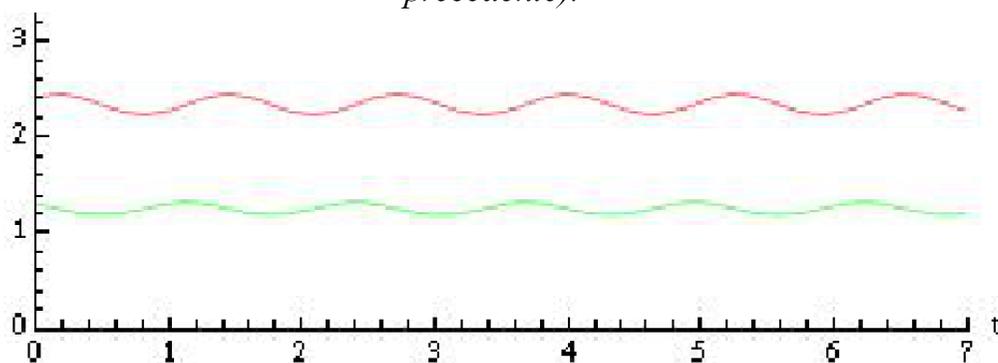


Illustrazione 4: Andamento delle popolazioni di prede (verde) - predatori (rosso) vs. tempo, nell'intorno del punto fisso centrale del grafico precedente.

Il massimo dei predatori è sempre sfasato in avanti rispetto al massimo delle prede.

Modello stocastico prede-predatori Wa-Tor

“pianeta” 2D con PBC (quindi toroidale), guerra tra due specie

Modello su reticolo quadrato : ogni sito e' vuoto o occupato da Preda o predatore.

Evoluzione con regole precise, e moto casuale tipo RW

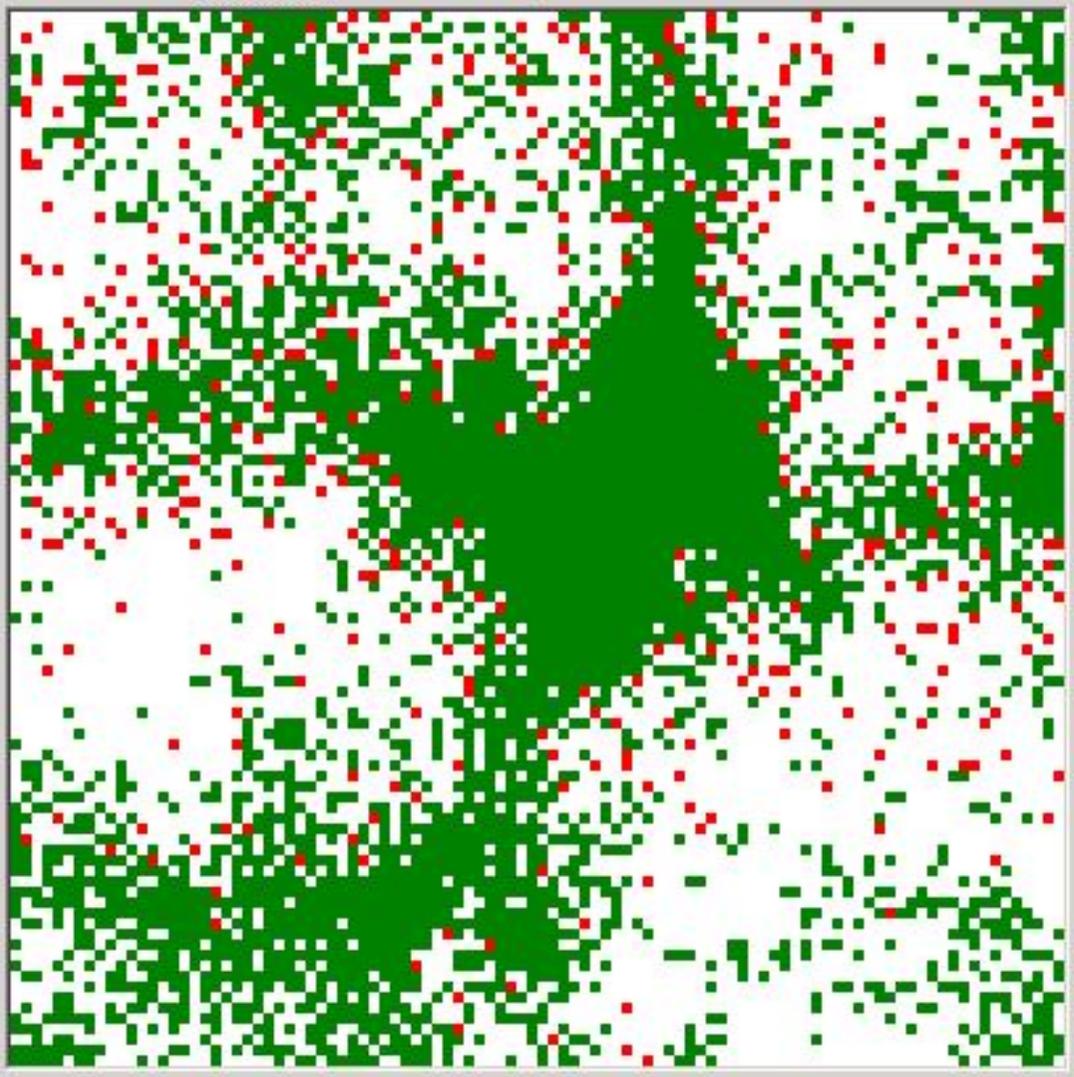
- Per simulare la riproduzione vengono introdotti due parametri, uno per ogni specie: se un individuo sopravvive per un numero di step Monte Carlo multiplo del tempo di riproduzione, esso si riproduce, se ha dello spazio libero intorno a se'.

Ad ogni pesce e ad ogni squalo viene quindi attribuita un'età casuale compresa tra zero ed il periodo di riproduzione.

Per simulare infine la predazione ogni squalo deve mangiare un pesce (trovandosi nel suo stesso punto del reticolo) entro un certo numero di step (altro parametro), altrimenti muore (viene rimosso dal reticolo)



Lotka-Volterra Wa-Tor Grafici Wa-Tor



Lato del reticolo
100

Lupi
n°: 800

Tempo di riproduzione
8

Tempo max. di digiuno
4

Pecore
n°: 2500

Tempo di riproduzione
8

durata della simulazione
4000

seme
1423

- Il modello di Lotka-Volterra è definito su uno spazio, assunto omogeneo, nel quale le posizioni dei singoli individui non vengono prese in considerazione: quest'approccio rappresenta quindi una descrizione di tipo campo-medio del sistema, nel quale non vengono descritte le correlazioni spaziali tra gli individui delle due specie.

-

Modello epidemiologico Reed-Frost



Modello SIR

(**S**uscettibili (o recettivi) / **I**nfetti / **R**esistenti (o immuni))

$$S(t) + I(t) + R(t) = \text{tot popolazione (costante)}$$

<http://www.quadernodiepidemiologia.it/epi/modelli/reed.htm>

<https://www.osc.edu/education/si/projects/epidemic>

Modello epidemiologico Reed-Frost



p probabilità (empirica...) per un dato individuo di avere un contatto efficiente, cioè che si infetti se era suscettibile; $q = 1 - p$

Formulazione stocastica:

$$\left\{ \begin{array}{l} I_{t+1} = \sum_{k=0}^{S_t} \mathcal{B}(1 - q^{I_t}) \\ S_{t+1} = S_t - I_{t+1} \end{array} \right.$$

$\mathcal{B}(x)$ variabile casuale che vale 1 con probabilità x , altrimenti 0

probabilità che almeno uno degli infetti al tempo t abbia un contatto efficiente

Con le condizioni iniziali di: popolazione totale, S_0 , I_0

Modello epidemiologico Reed-Frost



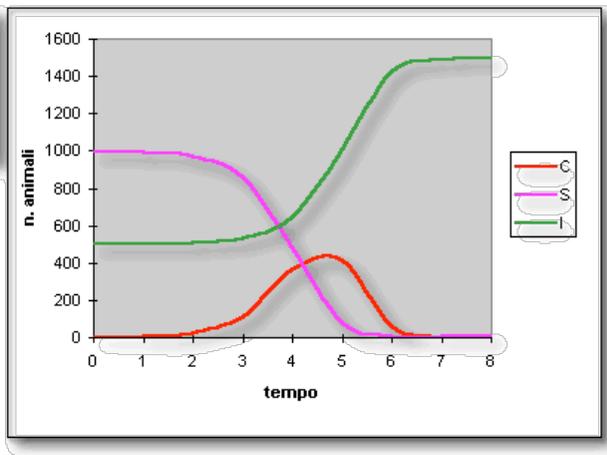
Il modello stocastico si può anche risolvere in modo deterministico sostituendo alle variabili casuali il loro valore medio:

$$\left\{ \begin{array}{l} I_{t+1} = \sum_{k=0}^{S_t} \mathcal{B}(1 - q^{I_t}) \\ S_{t+1} = S_t - I_{t+1} \end{array} \right. \Rightarrow \left\{ \begin{array}{l} I_{t+1} = S_t (1 - q^{I_t}) \\ S_{t+1} = S_t q^{I_t} \end{array} \right.$$

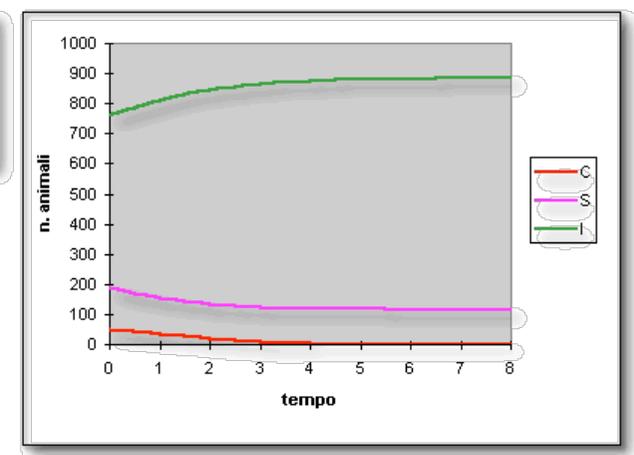
Riscrivendo: $S_{t+1} = S_t e^{-R_0}$ e confrontando con (*) per $I_t=1$ si ottiene:

$$R_0 = -\ln q$$

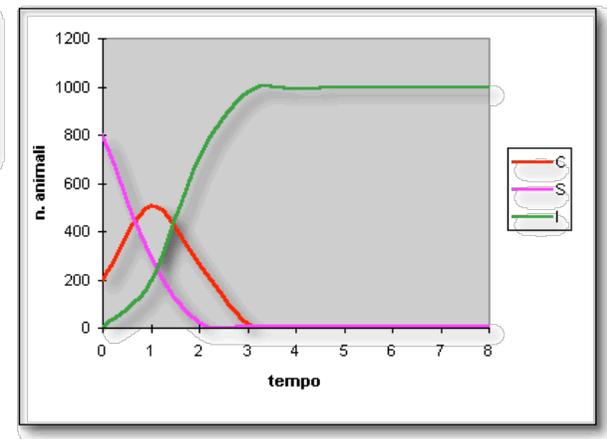
Casi	1		
N. di animali inizialmente recettivi	999		
N. di animali inizialmente immuni	500		
Dimensione della popolazione	1500		
Probabilità contatto efficiente	0.005		
Probabilità contatto NON efficiente	0.995		
Tempo	Casi	Suscett.	Immuni
t	C	S	I
0	1	999	500
1	5	994	501
2	25	969	506
3	112	857	531
4	369	488	643
5	411	77	1012
6	67	10	1423
7	3	7	1490
8	0	7	1493



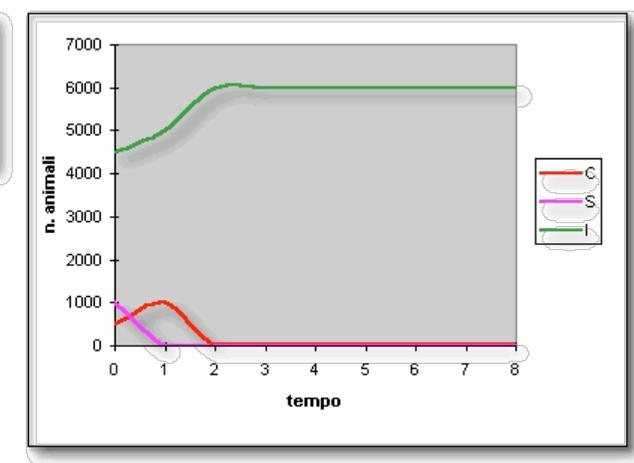
Casi	50		
N. di animali inizialmente recettivi	190		
N. di animali inizialmente immuni	760		
Dimensione della popolazione	1000		
Probabilità contatto efficiente	0.004		
Probabilità contatto NON efficiente	0.996		
Tempo	Casi	Suscett.	Immuni
t	C	S	I
0	50	190	760
1	35	155	810
2	20	135	845
3	10	125	865
4	5	120	875
5	2	117	880
6	1	116	883
7	1	116	884
8	0	115	884



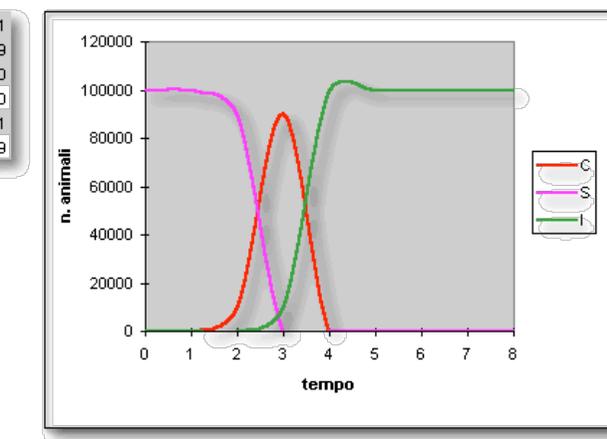
Casi	200		
N. di animali inizialmente recettivi	800		
N. di animali inizialmente immuni	0		
Dimensione della popolazione	1000		
Probabilità contatto efficiente	0.005		
Probabilità contatto NON efficiente	0.995		
Tempo	Casi	Suscett.	Immuni
t	C	S	I
0	200	800	0
1	506	294	200
2	270	23	706
3	17	6	977
4	0	5	994
5	0	5	995
6	0	5	995
7	0	5	995
8	0	5	995



Casi	500		
N. di animali inizialmente recettivi	1000		
N. di animali inizialmente immuni	4500		
Dimensione della popolazione	6000		
Probabilità contatto efficiente	0.06		
Probabilità contatto NON efficiente	0.94		
Tempo	Casi	Suscett.	Immuni
t	C	S	I
0	500	1000	4500
1	1000	0	5000
2	0	0	6000
3	0	0	6000
4	0	0	6000
5	0	0	6000
6	0	0	6000
7	0	0	6000
8	0	0	6000



Casi	1		
N. di animali inizialmente recettivi	99999		
N. di animali inizialmente immuni	0		
Dimensione della popolazione	100000		
Probabilità contatto efficiente	0.001		
Probabilità contatto NON efficiente	0.999		
Tempo	Casi	Suscett.	Immuni
t	C	S	I
0	1	99999	0
1	100	99899	1
2	9511	90388	101
3	90381	7	9612
4	7	0	99993
5	0	0	100000
6	0	0	100000
7	0	0	100000
8	0	0	100000



grande variabilità dei risultati in un modello epidemico attribuibile al solo caso!

Modello epidemiologico Reed-Frost



Limiti del modello:

- ipotesi ben poco realistica: che ogni individuo infetto incontri ogni altro individuo suscettibile nel periodo di tempo in cui e' infetto.
- unica modalità di trasmissione
- Periodo di incubazione+infezione fisso => determina Δt
- La popolazione è chiusa (nè ingressi nè uscite)
- Condizioni costanti durante tutta l'evoluzione epidemica

Modello epidemiologico Reed-Frost



Osservazioni varie:

- approccio stocastico importante e adeguato soprattutto nelle fasi iniziale/finale, dove i casi sono pochi e la casualità gioca un ruolo fondamentale
- Si può complicare a volontà, più facilmente di un modello deterministico

Altro approccio possibile (sempre stocastico):
modello su reticolo (si tiene conto di diversi modi di diffusione, e della mobilità degli individui)

Buon divertimento!