Passeggiata Zombi

Angelo Greco

Giugno 2025

1 Introduzione

Il "**problema degli zombi**" è un quesito di probabilità ispirato a uno scontro fittizio tra due armate: un gruppo di esseri umani e un gruppo di zombi. Ogni combattente umano affronta uno zombi in duello, con probabilità di successo $\frac{1}{2}$. Se vince, sfida il prossimo zombi; se perde, l'umano muore e viene convertito in un nuovo zombi.

Lo scopo di questo progetto è di studiare questa dinamica mediante *modelli probabilistici*, in particolare tramite l'uso di passeggiate aleatorie, e stimare quanti combattenti umani siano necessari per avere almeno l'80% di probabilità di eliminare tutti gli zombi.

2 Teoria

In questa sezione, verranno presentati degli accenni teorici agli argomenti rilevanti per tale problema. Sia $(\Omega, \mathcal{F}, \mathcal{P})$ uno spazio di probabilità fissato.

2.1 Tempo di Arresto e Tempo di Uscita

Sia $(\xi_n)_{n\in\mathbb{N}}$ un processo stocastico a valori in uno spazio discreto. Una variabile aleatoria $\tau : \Omega \to \mathbb{N} \cup \{\infty\}$ è detta **tempo di arresto** (rispetto alla filtrazione $(\mathcal{F}_n^\xi)_{n\in\mathbb{N}}$) se:

$$\{\tau \leq n\} \in \mathcal{F}_n^\xi, \quad \forall n \in \mathbb{N}$$

ovvero, con l'informazione disponibile fino a tempo n, si può determinare se un evento che accade a tempo τ sia avvenuto o meno.

Una variante del tempo d'arresto è il **tempo di prima uscita** da un insieme $A \subseteq \mathbb{R}$ per un processo stocastico ξ_n, definito come:

$$\tau_A := \inf\{n \geq 0 : \xi_n \notin A\}$$

ovvero, il tempo di prima uscita τ_A è il primo istante n in cui il processo ξ_n non appartiene più all'insieme A.

2.2 Martingala e Passeggiata Aleatoria

Sia $(X_n)_{n\in\mathbb{N}_0}$ un processo stocastico a valori reali, adattato a una filtrazione $(\mathcal{F}_n^\xi)_{n\in\mathbb{N}_0}$. Diciamo che X è una:

- **sub-martingala** se:
  1. $X_n \in L^1 \quad \forall n \in \mathbb{N}_0 \qquad$ (quindi $E[X_n]$ definito),
  2. $E[X_{n+1} \mid \mathcal{F}_n^\xi] \geq X_n \quad \forall n \in \mathbb{N}_0$;
- **super-martingala** se vale la stessa definizione, ma con:

$$E[X_{n+1} \mid \mathcal{F}_n^\xi] \leq X_n \quad \forall n \in \mathbb{N}_0;$$

- **martingala** se invece vale:

$$E[X_{n+1} \mid \mathcal{F}_n^\xi] = X_n \quad \forall n \in \mathbb{N}_0.$$

Sia ora $(\xi_n)_{n\in\mathbb{N}}$ una successione di variabili aleatorie reali. Definiamo la filtrazione associata come:

$$\mathcal{F}_0^\xi := \{\emptyset, \Omega\}, \quad \mathcal{F}_n^\xi := \sigma(\xi_1, \ldots, \xi_n) \quad \forall n \in \mathbb{N}_0.$$

Dato un valore iniziale $X_0 \in \mathbb{R}$, costruiamo un processo $(X_n)_{n\in\mathbb{N}_0}$ definito ricorsivamente come:

$$X_{n+1} := X_n + \xi_{n+1}, \quad \forall n \in \mathbb{N}.$$

Tale processo rappresenta una **passeggiata aleatoria** reale, con punto di partenza X_0 e incrementi dati dalla successione ξ_n.

Se gli incrementi ξ_n sono indipendenti e con *valore medio nullo*, il processo X è una *martingala* rispetto a $\mathcal{F}_n^\xi$. Infatti, si ha:

$$E[X_{n+1} \mid \mathcal{F}_n^\xi] = E[X_n + \xi_{n+1} \mid \mathcal{F}_n^\xi] = X_n + E[\xi_{n+1} \mid \mathcal{F}_n^\xi] = X_n + 0 = X_n.$$

2.3 Disuguaglianza massimale di Kolmogorov

Siano Y_i, con $i \in \mathbb{N}$, variabili aleatorie indipendenti, con *media nulla* e *varianza finita*. Si definisce:

$$S_0 = 0, \quad S_n := \sum_{i=1}^n Y_i \quad \text{per ogni } n \geq 1.$$

Allora, per ogni $x > 0$, vale la seguente disuguaglianza:

$$\mathcal{P}\left(\max_{1\leq k\leq n} |S_k| \geq x\right) \leq \frac{1}{x^2}\,\mathrm{Var}(S_n) = \frac{1}{x^2}\sum_{i=1}^n \mathrm{Var}(Y_i).$$

2.4 Moto browniano e principio d'invarianza di Donsker

Il **moto browniano** è un processo stocastico continuo $(B_t)_{t\geq 0}$, con incrementi *indipendenti* e *distribuiti normalmente*: per ogni $0 \leq s < t$ si ha che $B_t - B_s \sim \mathcal{N}(0, t-s)$.

Il **principio d'invarianza di Donsker** afferma che una passeggiata aleatoria normalizzata opportunamente, a passi *i.i.d.* (indipendenti e identicamente distribuiti), converge in distribuzione a un moto browniano. In particolare, se S_n è una passeggiata aleatoria con incrementi a *media nulla* e *varianza finita*, allora:

$$\left(\frac{S_{\lfloor n^2 t \rfloor}}{n}\right)_{t\geq 0} \xrightarrow[n\to\infty]{\mathcal{D}} B_t$$

dove con l'operatore "*floor*" $\lfloor x \rfloor$ prendiamo il più grande intero minore o uguale a x.

3 Definizione del problema

Definiamo una collezione di variabili aleatorie X_i, ciascuna rappresentante il numero di zombi uccisi dall'i-esimo combattente umano prima di morire. Queste variabili seguono una *distribuzione geometrica a supporto zero*, cioè $X_i \sim \mathrm{Geo}_0(p)$ con $p = \frac{1}{2}$:

$$\mathcal{P}(X_i = k) = (1-p)^k p = \left(\frac{1}{2}\right)^{k+1}, \quad k \in \mathbb{N}_0.$$

Inoltre, si può assumere che tali variabili aleatorie $(X_i)_{i\in\mathbb{N}}$ siano indipendenti (poiché ogni combattente effettua i propri duelli individualmente).

3.1 La variabile Y_i

Dato che, quando un combattente muore, diventa uno zombi, definiamo:

$$Y_i := X_i - 1$$

che rappresenta il contributo netto dell'i-esimo umano alla riduzione della popolazione zombi. Dimostriamo che tale variabile ha *media nulla* e *varianza pari a 2*.

Media:

$$E[Y_i] = E[X_i - 1] = E[X_i] - 1.$$

La media di una distribuzione geometrica con supporto zero è:

$$E[X_i] = \frac{1-p}{p} = \frac{1-\frac{1}{2}}{\frac{1}{2}} = 1.$$

Quindi:

$$E[Y_i] = 1 - 1 = 0.$$

Varianza:

$$\mathrm{Var}(Y_i) = \mathrm{Var}(X_i - 1) = \mathrm{Var}(X_i) \quad \text{(traslare non cambia la varianza).}$$

La varianza di una distribuzione geometrica con supporto zero è:

$$\mathrm{Var}(X_i) = \frac{1-p}{p^2} = \frac{1-\frac{1}{2}}{\left(\frac{1}{2}\right)^2} = \frac{\frac{1}{2}}{\frac{1}{4}} = 2.$$

Conclusione:

$$E[Y_i] = 0, \quad \mathrm{Var}(Y_i) = 2.$$

3.2 La passeggiata aleatoria zombi

Poniamo:

$$S_n := \sum_{i=1}^{n} Y_i.$$

Allora S_n è una passeggiata aleatoria, che inizia da $S_0 = 0$. Essa traccia il numero di zombi *in meno* dopo la morte dell'n-esimo combattente.

Definiamo ora il tempo di arresto come:

$$T_m := \inf\{n \geq 0 : S_n \geq m\}.$$

Con una popolazione iniziale di zombi pari a M, i combattenti umani vincono appena $S_n \geq M - 1$, ovvero all'istante aleatorio T_{M-1}: quando un combattente uccide l'ultimo zombi, non serve considerare l'eventualità in cui muore e si trasforma in uno zombi.

3.3 Stima probabilistica del successo degli umani

Indichiamo con N la taglia dell'armata dei combattenti umani. La popolazione zombi sarà eliminata se:

$$T_{M-1} \leq N.$$

Il problema è dunque determinare il valore critico di $N = N(M)$ tale che:

$$\mathcal{P}(T_{M-1} \leq N) \approx 0.8\,.$$

Osserviamo che l'evento $\{T_{M-1} \leq N\}$ è equivalente all'esistenza di un indice $n \leq N$ per cui $S_n \geq M - 1$. In simboli:

$$\{T_{M-1} \leq N\} = \{\exists n \leq N \text{ t.c. } S_n \geq M-1\} = \left\{\max_{1 \leq k \leq N} S_k \geq M - 1\right\}.$$

Possiamo ora applicare la **disuguaglianza massimale di Kolmogorov**, che afferma che per una somma parziale di variabili aleatorie indipendenti con *media nulla* e *varianza finita*:

$$\mathcal{P}\left(\max_{1 \leq k \leq N} |S_k| \geq M - 1\right) \leq \frac{\mathrm{Var}(S_N)}{(M-1)^2}\,.$$

Nel nostro caso, essendo $Y_i = X_i - 1$ indipendenti e con varianza $\mathrm{Var}(Y_i) = 2$, otteniamo:

$$\mathrm{Var}(S_N) = \sum_{i=1}^{N} \mathrm{Var}(Y_i) = 2N.$$

Quindi:

$$\mathcal{P}(T_{M-1} \leq N) = \mathcal{P}\left(\max_{1\leq k\leq N} S_k \geq M-1\right) \leq \mathcal{P}\left(\max_{1\leq k\leq N} |S_k| \geq M-1\right) \leq \frac{2N}{(M-1)^2}.$$

Per ottenere una probabilità *almeno* pari all'80% di successo, la disuguaglianza di Kolmogorov richiede che:

$$\frac{2N}{(M-1)^2} \geq 0.8\,.$$

Isolando N, otteniamo la seguente condizione sul numero minimo di combattenti umani:

$$N \geq \frac{0.8 \cdot (M-1)^2}{2} = 0.4 \cdot (M-1)^2.$$

Ne segue che, *per vincere con probabilità 80%, il numero di combattenti deve essere* ***almeno*** $0.4 \cdot (M-1)^2$.

Esempio pratico: Supponiamo di avere $M = 100$ zombi. Applichiamo la stima ottenuta:

$$N \geq 0.4 \cdot (M-1)^2 = 0.4 \cdot 99^2 = 0.4 \cdot 9801 = 3\,920.4\,.$$

Dunque, secondo la disuguaglianza di Kolmogorov, servono *almeno* 3 920 combattenti per avere una probabilità di vittoria dell'80%.
Tuttavia, non ci accontentiamo di un limite inferiore: le simulazioni svolte suggeriscono che questa stima è fin troppo ottimista, e il numero effettivo di umani necessari per vincere con l'80% di probabilità è molto più alto.

4 Simulazioni

Per simulare gli scontri inerenti al progetto, sono state implementate delle funzioni in Python, i cui risultati sono stati racchiusi in grafi.

4.1 Variabile aleatoria X_i

La variabile aleatoria X_i è generata tramite una funzione della libreria `numpy` che genera un valore con distribuzione geometrica usando una probabilità di 0.5 (ovvero $\frac{1}{2}$), a cui sottraiamo 1 per supportare anche lo zero:

```
def X():
    return np.random.geometric(0.5) - 1
```

Questa implementazione è stata preferita ad altre per la sua efficienza. Una versione artigianale può consistere in un contatore inizializzato a zero, che viene aumentato se un valore casuale generato in modo uniforme tra 0 e 1 è minore di 0.5, e ritornato altrimenti.

4.2 Passeggiata aleatoria S_n

La passeggiata aleatoria zombi S_n è simulata tramite una funzione che prende in input il numero N di combattenti umani e il numero M di zombi iniziali, e, partendo da zero, usa come valore di ogni passo successivo (rappresentato dalla morte di un combattente) la variabile aleatoria Y_i (ovvero $X_i - 1$).

```
def PasseggiataZombi(N, M):
    S = [0]
    for i in range(N):
        if S[i] < M - 1:
            S.append(min(S[i] + X() - 1, M))
        else:
            S.append(M)
    return S
```

In questa implementazione, una volta che la passeggiata raggiunge $M - 1$ (cioè una volta che sono stati sconfitti tutti gli zombi), S_n mantiene il valore M per tutti gli $n \geq M - 1$, rappresentando l'assenza di futuri scontri.

La seguente figura mostra la simulazione di una passeggiata aleatoria S_n per un singolo esperimento:

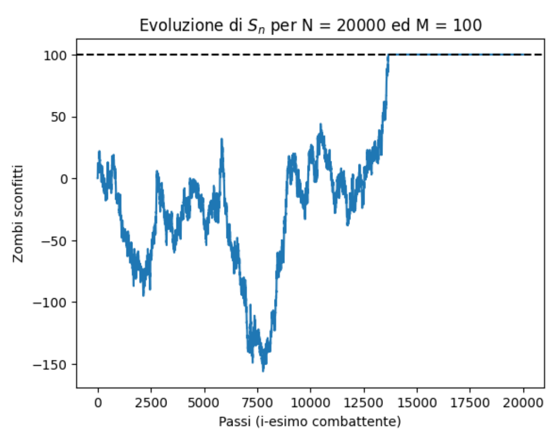

Figure 1: Evoluzione di una passeggiata zombi con $M = 100$ ed $N = 20\,000$. Se viene raggiunta la quota di *almeno* 99, vuol dire che gli zombi sono stati sconfitti ed S_n viene impostata ad $M = 100$.

Per evidenziare la variabilità dell'esperimento, si riportano di seguito quattro simulazioni indipendenti della passeggiata S_n con gli stessi parametri:

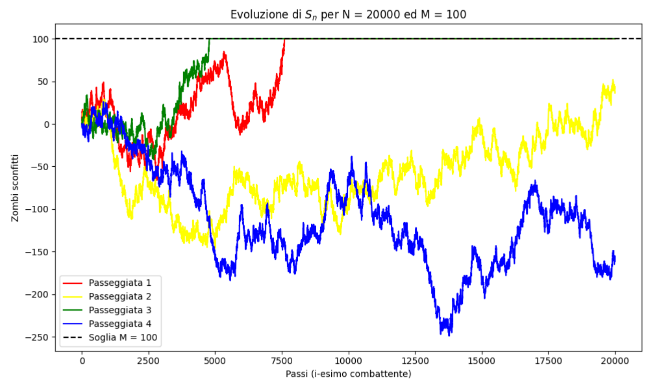

Figure 2: Quattro passeggiate aleatorie zombi indipendenti tra loro, con $M = 100$ e $N = 20\,000$. Si osserva che, in alcuni casi (passeggiata 2 e 4), i combattenti umani non sono stati in grado di sconfiggere la popolazione zombi.

4.3 Simulazione del tempo di arresto T_{M-1}

Per stimare la probabilità che un certo numero N di combattenti umani sconfiggano tutti gli zombi, abbiamo introdotto una funzione che misura il numero di combattenti richiesti per ridurre la popolazione zombi da M a 0 (ossia raggiungere $S_n \geq M - 1$). Questo è fatto prendendo in input il numero M di zombi iniziali, assieme ad una *soglia* massima per evitare di svolgere passeggiate troppo lunghe.

```
def T(M, soglia):
    i, S = 0, 0
    while S < M - 1 and i < soglia:
        S += X() - 1
        i += 1
    return i
```

Abbiamo eseguito $k = 1000$ simulazioni con $M = 100$ e $soglia = 10 \cdot M^2 + 1$, e con i valori ritornati è stata costruita la funzione di ripartizione *approssimata* di T_{M-1}/M^2 (quindi normalizziamo per avere la disuguaglianza $T_{M-1}/M^2 \leq x$ al posto di $T_{M-1} \leq xM^2$). Tale funzione è:

$$F_k(x) := \frac{1}{k} \sum_{j=1}^{k} \mathbf{1}_{\{T^{(j)}/M^2 \leq x\}}.$$

Essa rappresenta una stima della probabilità che il numero normalizzato di combattenti necessari sia inferiore o uguale a x, ovvero:

$$F_k(x) \approx \mathcal{P}\left(\frac{T_{M-1}}{M^2} \leq x\right).$$

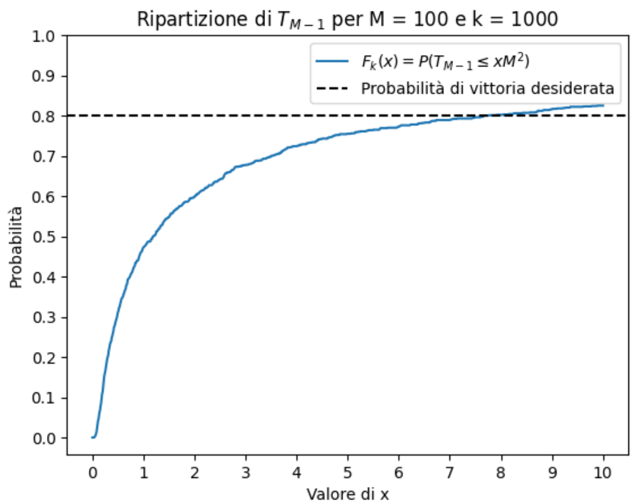

Figure 3: Funzione di ripartizione *approssimata* $F_k(x)$ ottenuta tramite $k = 1000$ simulazioni per $M = 100$, con *soglia* $= 10 \cdot M^2 + 1$. L'aggiunta di "+1" evita la presenza di artefatti nell'estremo destro del grafico causati dalle simulazioni troncate.

Osservando il grafico si nota che la probabilità di successo raggiunge l'80% quando $x \approx 8$, da cui può essere effettuata la seguente stima:

$$N \approx 8 \cdot M^2.$$

Servirebbero quindi $8 \cdot 100^2 = 80\,000$ combattenti umani per avere l'80% di probabilità di sconfiggere 100 zombi!

5 Approfondimenti sulla stima ottimale

Aumentando il numero di simulazioni effettuate ($k = 100\,000$) possiamo determinare con maggiore precisione il valore ottimale di x per cui $P(T_{M-1}/M^2 \leq x) \approx 0.8$. Con questo metodo, si ottiene un'approssimazione pari a:

$$x \approx 7.795.$$

Questa stima è di certo più accurata, ma pur sempre empirica; per raffinarla ulteriormente si può effettuare un'analisi asintotica del processo.

5.1 Approssimazione asintotica della distribuzione di T_{M-1}

Per grandi valori di M, la variabile T_{M-1} può essere approssimata tramite un modello continuo, grazie al **principio d'invarianza di Donsker**. Considerando la definizione della passeggiata aleatoria $S_n := \sum_{i=1}^{n}(X_i - 1)$, con incrementi *i.i.d.*, *media nulla* e *varianza finita*, si ha la seguente convergenza in distribuzione:

$$\left(\frac{S_{\lfloor M^2 t\rfloor}}{M}\right)_{t\geq 0} \xrightarrow[M\to\infty]{\mathcal{D}} (B_t)_{t\geq 0}$$

dove $(B_t)_{t\geq 0}$ è un **moto browniano**.

Si osserva innanzitutto che la condizione di arresto $S_n \geq M-1$ può essere riscritta dividendo entrambi i membri per M:

$$S_n \geq M-1 \quad \Longleftrightarrow \quad \frac{S_n}{M} \geq \frac{M-1}{M}.$$

Per avvicinarci a un'analisi asintotica continua, introduciamo una nuova variabile di tempo:

$$t := \frac{n}{M^2} \quad \Longleftrightarrow \quad n = \lfloor M^2 t \rfloor,$$

così da poter esprimere la passeggiata aleatoria S_n come funzione del tempo continuo t, almeno approssimativamente:

$$\frac{S_n}{M} \approx \frac{S_{\lfloor M^2 t\rfloor}}{M}.$$

Ricordando la definizione del tempo di arresto:

$$T_{M-1} := \inf\{n \geq 0 : S_n \geq M-1\},$$

la riscriviamo in termini del tempo continuo t come:

$$\frac{T_{M-1}}{M^2} = \inf\left\{t \geq 0 : \frac{S_{\lfloor M^2 t\rfloor}}{M} \geq \frac{M-1}{M}\right\}.$$

Poiché $\frac{M-1}{M} \to 1$ per $M \to \infty$, il momento in cui il processo normalizzato supera $\frac{M-1}{M}$ è sempre più vicino al momento in cui il moto browniano supera 1. Quindi il tempo del processo normalizzato converge al tempo di uscita del moto browniano dall'intervallo $[0,1)$:

$$\frac{T_{M-1}}{M^2} \xrightarrow[M\to\infty]{\mathcal{D}} \inf\{t \geq 0 : B_t \geq 1\} := \tau.$$

La funzione di ripartizione di τ è pari a:

$$\mathbb{P}(\tau \leq x) = 2\left(1 - \Phi\left(\frac{1}{\sqrt{2x}}\right)\right)$$

dove Φ è la funzione di ripartizione della distribuzione normale $\mathcal{N}(0,1)$. Tale formula fornisce l'approssimazione limite della distribuzione di T_{M-1}/M^2.

Il valore "teorico" di x per cui questa probabilità è 0.8 può essere quindi calcolato risolvendo l'equazione:

$$2\left(1-\Phi\left(\frac{1}{\sqrt{2x}}\right)\right)=0.8\,.$$

Per risolverla usiamo il *metodo di Brent*[1], implementato tramite libreria `scipy`:

```
def funzione_limite(x):
    return 2 * (1 - norm.cdf(1 / np.sqrt(2 * x)))

def trova_x_teorico(perc_desiderata=0.8):
    def f(x):
        return funzione_limite(x) - perc_desiderata
    sol = root_scalar(f, bracket=[1e-3, 20], method='brentq')
    return sol.root
```

Ottenendo il risultato pari a:

$$x \approx 7.79$$

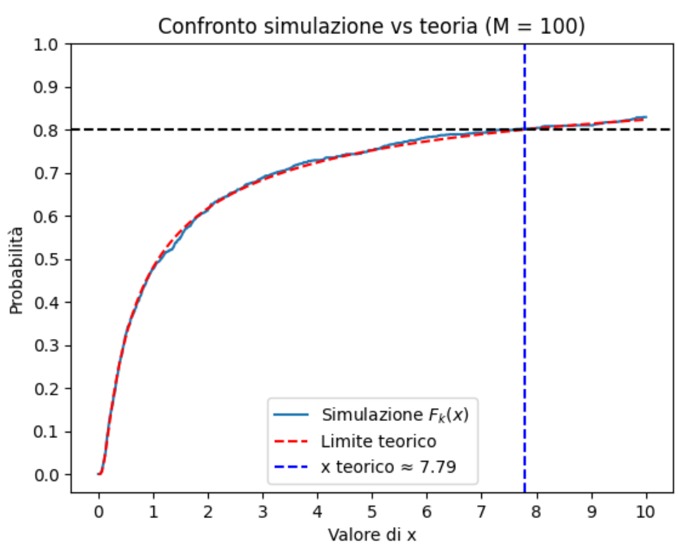

Figure 4: Funzione di ripartizione $2(1-\Phi(1/\sqrt{2x}))$ (in rosso) assieme a una simulazione (in azzurro), con il valore $x \approx 7.79$ evidenziato.

[1] Documentazione: `https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.root_scalar.html`

Pertanto, possiamo concludere questo approfondimento affermando che:

Per vincere con probabilità 80%, servono circa $7.79 \cdot M^2$ combattenti.

6 Conclusioni

Abbiamo modellato la dinamica del conflitto tra umani e zombi come una passeggiata aleatoria, basata su variabili geometriche. Le simulazioni effettuate hanno permesso di approssimare un valore ottimale per il numero di combattenti umani in funzione del numero di zombi iniziali $N(M)$, in seguito approfondito tramite lo studio del comportamento limite del processo: il valore ottimale trovato è $N \approx 7.79M^2$.

Questo esercizio dimostra come anche fenomeni all'apparenza semplici, come una passeggiata aleatoria, possano dare origine a comportamenti complessi e non immediatamente prevedibili: piccole variazioni nel rapporto tra umani e zombi possono alterare drasticamente le probabilità di sopravvivenza. Il lavoro effettuato mostra sia la potenza degli strumenti probabilistici e computazionali nella rappresentazione e simulazione di sistemi complessi, sia la loro capacità di facilitarne la comprensione, attraverso la creazione di dimostrazioni ed esempi.