

Modellizzazione dei Sistemi Logistici

Algoritmi genetici

Corso di Studi in Ingegneria Gestionale

a.a. 2020-2021

giuseppe.bruno@unina.it

Docente: Prof. Giuseppe Bruno

Argomenti della lezione

- **Introduzione**
- **Caratteristiche della procedura**
- **I passi della procedura**
- **Considerazioni generali**

Massimo globale e locale

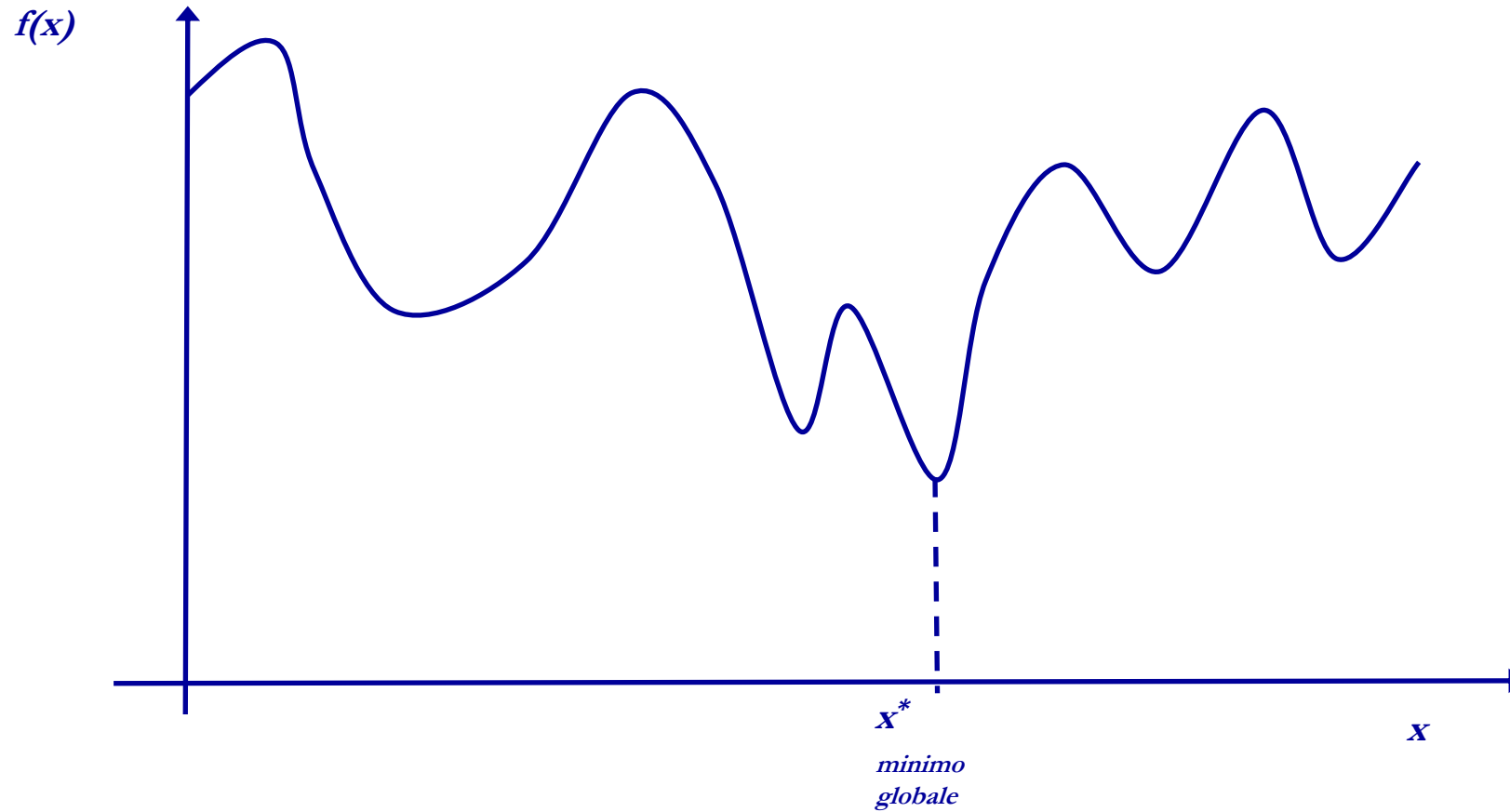
In presenza di problemi complessi dal punto di vista computazionale (NP-hard) nelle applicazioni è necessario ricorrere a **tecniche euristiche di risoluzione**.

Le tecniche costruttive sono orientate alla individuazione (costruzione) di una soluzione ammissibile.

Le tecniche migliorative o di ricerca locale partono da una soluzione ammissibile e puntano a migliorarla attraverso la applicazione iterativa di una mossa. Esse convergono in un **ottimo locale**.

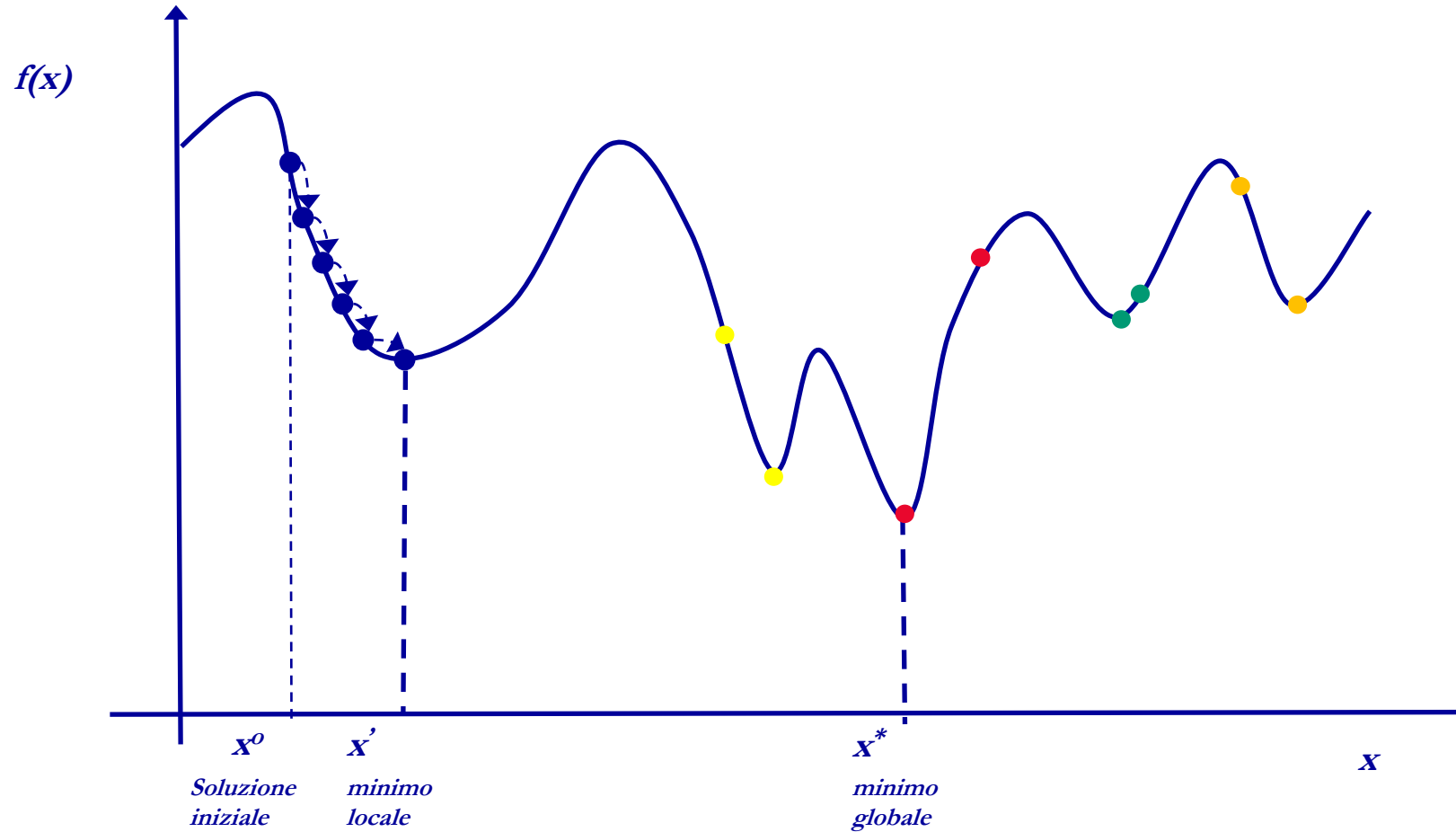
Per migliorare la qualità delle soluzioni individuabili con queste tecniche è possibile ricorrere a tecniche cosiddette **metaeuristiche**

Massimo globale e locale



Massimo globale e locale

Convergenza di una procedura di ricerca locale in un minimo locale



Il metodo della roulette

Detto anche metodo Montecarlo, prende il nome dal quartiere del Principato di Monaco, che ospita il famoso e antico (1856) casinò.

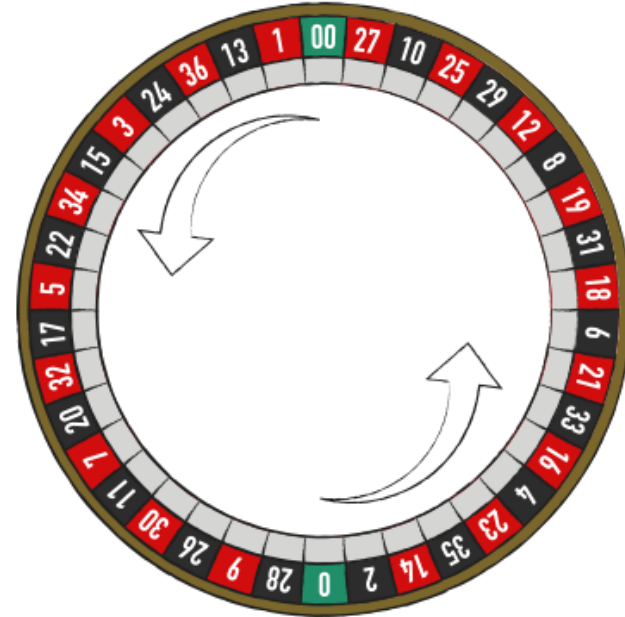


Il metodo della roulette

Il gioco si basa sul fatto che, a partire da condizioni iniziali casuali, una pallina lanciata in direzione opposta al senso di rotazione (rulèt: ruotare), possa alla fine posizionarsi, con la stessa probabilità in uno dei settori (37 o 38), di uguale ampiezza, in cui è divisa la ruota di antico (1856) casinò.



Roulette francese



Roulette americana

Il metodo della roulette

Il metodo consente di simulare eventi probabilistici caratterizzati da una predefinita funzione discreta di probabilità.

Si considerino n possibili eventi ciascuno caratterizzato da una probabilità p_i (ovviamente $\sum_{i=1,n} p_i = 1$).

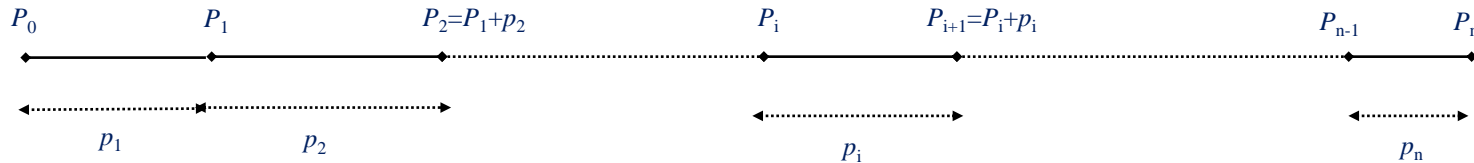
Per simulare il lancio di una pallina sulla roulette:

- Si calcola per ciascun elemento i il *valore cumulato di probabilità* $P_i = \sum_{k=1,i} p_k$ o attraverso la *formula ricorsiva* $P_i = P_{i-1} + p_i$, ponendo $P_0 = 0$;
- Si genera un numero casuale a tra 0 e 1 ($0 \leq a \leq 1$);
- Si seleziona l'elemento j tale che $P_{j-1} \leq a < P_j$

Il metodo della roulette

Per simulare il lancio di una pallina sulla roulette:

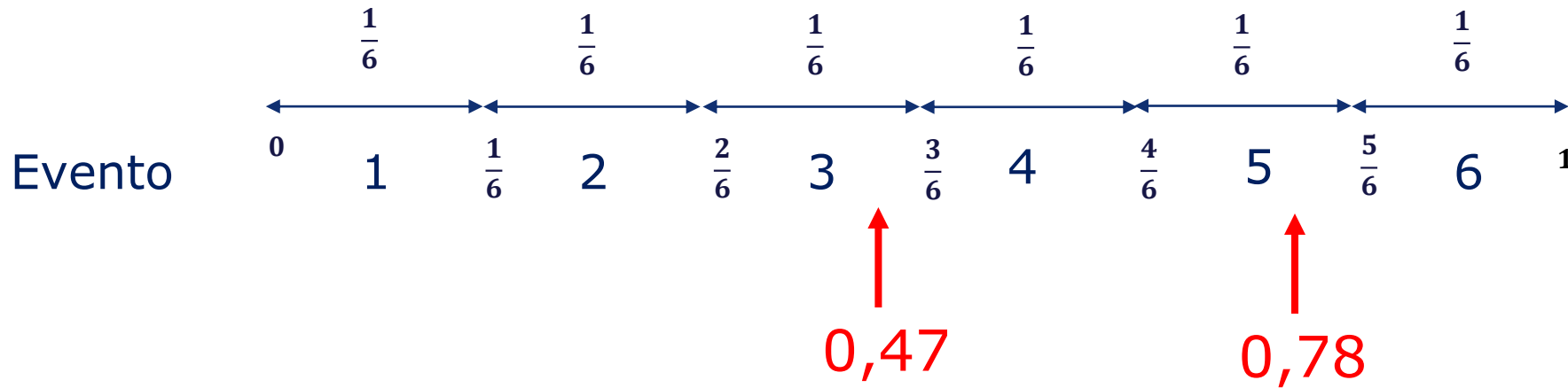
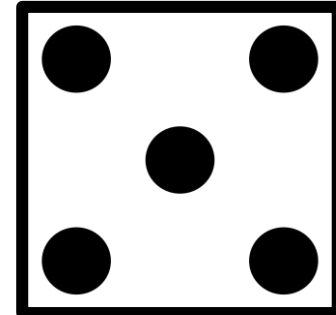
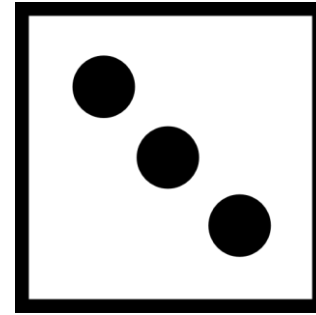
- Si calcola per ciascun elemento i il *valore cumulato di probabilità* $P_i = \sum_{k=1,i} p_k$ o attraverso la *formula ricorsiva* $P_i = P_{i-1} + p_i$, ponendo $P_0 = 0$;
- Si genera un numero casuale a tra 0 e 1 ($0 \leq a \leq 1$);
- Si seleziona l'elemento j tale che $P_{j-1} \leq a < P_j$



Il metodo della roulette

Esempio: si simuli il lancio di due dadi
 $n=6$ eventi tutti di pari probabilità

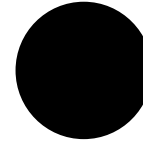
$$p_i = \frac{1}{6} \quad i = 1..6$$



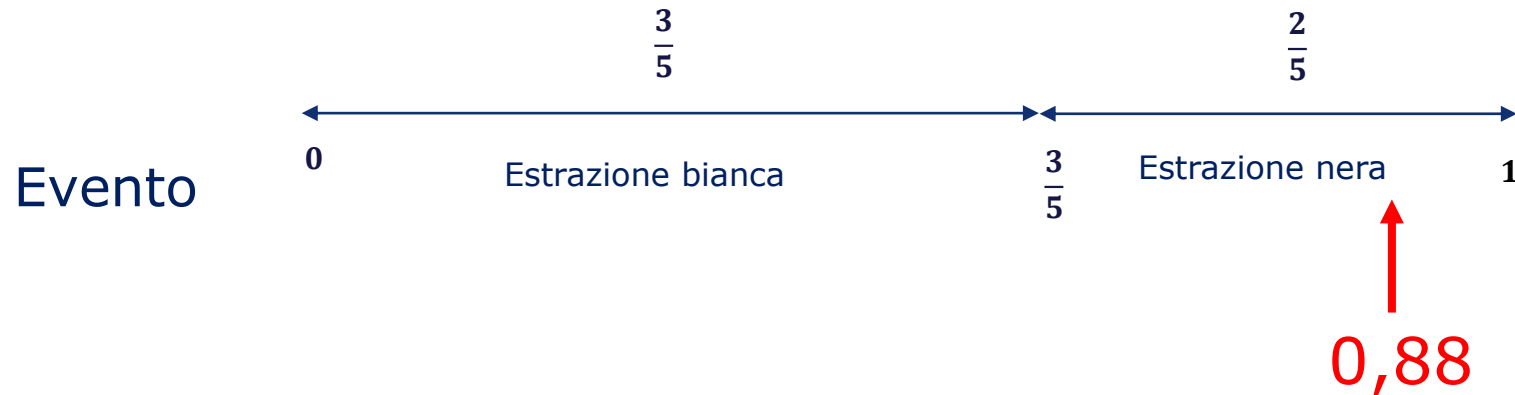
Generazione due numeri casuali: 0,47 0,78

Il metodo della roulette

Esempio: si l'estrazione di una pallina da un'urna che ne contiene 3 bianche e 2 nere
 $n=2$ di probabilità



$$p_1 = \frac{3}{5} \quad p_2 = \frac{2}{5}$$

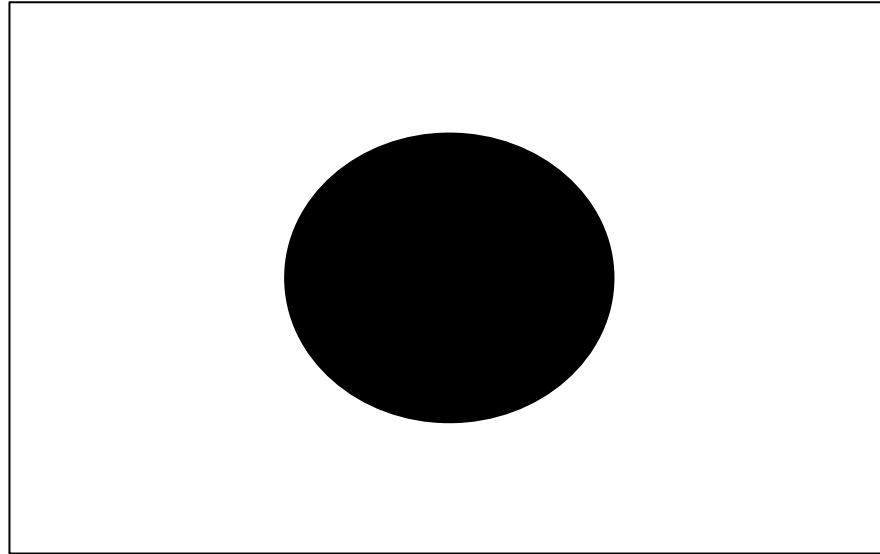


Generazione numero casuale: 0,88

Il metodo della roulette

Ogni linguaggio ha una funzione generatrice di numeri (pseudo)casuali

In Excel esistono le funzioni `casuale()` e `casuale.tra(primo,secondo)`



Tecniche euristiche

La qualità della soluzione ottenibile da una tecnica di ricerca locale dipende dalla soluzione di innesto e dalla tipologia di mossa.

Per superare la “trappola” della convergenza in un minimo locale, in corrispondenza di esso, bisognerebbe individuare una procedura che consenta di accettare soluzioni peggiorative.

Algoritmi genetici

Gli **Algoritmi Genetici** sono una **tecnica metaeuristica naturale** che si basa sull'analogia con i meccanismi di selezione naturale in campo genetico.

L'idea di base è quella di considerare una "popolazione" di soluzioni che evolve in accordo con un meccanismo di selezione in modo da produrre soluzioni con buoni valori della funzione obiettivo.

Algoritmi genetici

L'evoluzione di una popolazione è legata al processo di riproduzione.

- Durante la loro vita gli individui (**genitori**) si accoppiano producendo nuovi individui (**figli**) il cui patrimonio genetico è una combinazione di quello dei genitori.
- I figli subiscono mutazioni rispetto al patrimonio genetico ereditato dai genitori per effetto della vita di relazione e delle influenze dell'ambiente.

La legge di selezione naturale si basa sul principio che, tra gli individui generati, **hanno maggiori probabilità di sopravvivere quelli che possiedono una fitness migliore.**

Algoritmi genetici

Le caratteristiche di un individuo sono determinate dai **geni** che custodiscono informazioni relative ad una specificità

L'insieme dei geni è detto **genotipo**.

Ciascun gene può assumere **alleli** diversi che producono differenze delle caratteristiche associate a quel gene.

Con il termine **fitness** ("forma fisica"), si intende la capacità di adattarsi ad un determinato ambiente.

L'evoluzione è un processo che altera, di generazione in generazione, le caratteristiche genetiche degli organismi in modo che possano adattarsi meglio al proprio ambiente. In altri termini, **il processo di evoluzione è la progressiva selezione di individui con elevata fitness.**

Algoritmi genetici

Corrispondenza evoluzione genetica - algoritmi genetici

Evoluzione genetica	Algoritmo genetico
<i>Gene</i>	<i>Variabile decisionale</i>
<i>Allele</i>	<i>Valore variabile decisionale</i>
<i>Cromosoma</i>	<i>Insieme variabili decisionali</i>
<i>Genotipo</i>	<i>Soluzione del problema</i>
<i>Fitness</i>	<i>Funzione obiettivo</i>
<i>Accoppiamento</i>	<i>Crossover</i>
<i>Influenza dell'ambiente</i>	<i>Altri operatori genetici</i>
<i>Selezione naturale</i>	<i>Algoritmo</i>

Algoritmi genetici

Gli *Algoritmi Genetici* simulano il processo di evoluzione partendo da una **popolazione iniziale** ed applicando ad essa i cosiddetti **operatori genetici**.

Considerando i problemi di ottimizzazione, ad un **gene** corrisponde una variabile decisionale e ad un **allele** il valore ad essa associato.

Ad un **cromosoma** corrisponde un insieme di variabili decisionali mentre il **genotipo**, ossia l'insieme dei valori assunti dai geni, è una possibile soluzione del problema.

La **fitness** associata ad ogni genotipo è il valore di funzione obiettivo.

I passi della procedura

1. Codifica delle soluzioni

Si rappresenta una soluzione in termini di stringa di variabili

2. Generazione popolazione iniziale e valutazione fitness

Si genera un insieme di possibili soluzioni che forma la popolazione iniziale e si associa a ciascuna soluzione un valore di fitness.

3. Selezione

Si selezionano coppie di soluzioni della popolazione alle quali applicare gli operatori genetici (crossover, mutazione...).

4. Generazione di nuove soluzioni

Si applicano gli operatori genetici alla coppia di soluzioni selezionate al fine di produrre nuove soluzioni.

5. Sostituzione di elementi della popolazione

Si sostituiscono le soluzioni prodotte a soluzioni presenti nella popolazione

6. Criterio di arresto

Si applicano iterativamente i passi 3, 4 e 5 fin quando non si verifichi una condizione di arresto.

I passi della procedura

Schema logico di funzionamento della procedura



I passi della procedura

1 - Codifica delle soluzioni

Si possono adottare due tipologie di codifica:

- **Codifica binaria**

Per problemi a combinazione (es: zaino)

- **Codifica a permutazione**

Per problemi a permutazione (es: TSP)

0	0	1	0	1	1	0	0
---	---	---	---	---	---	---	---

Esempio codifica binaria

8	3	6	4	5	1	7	2
---	---	---	---	---	---	---	---

Esempio codifica a permutazione

I passi della procedura

Generazione popolazione iniziale e valutazione della fitness

Si genera una popolazione iniziale di n individui e, per ciascuno di essi si calcola la fitness.

La popolazione può essere generata in maniera random o attraverso l'uso di algoritmi (costruttivi e/o migliorativi).

1	1	0	0	1	1	0	0	1
2	0	1	0	0	1	0	1	0
3	0	1	0	0	0	1	1	1
4	1	0	1	1	0	1	0	1
5	1	0	0	1	1	0	1	0
...	1	0	1	0	0	1	1	0
n-1	0	0	0	1	0	0	1	0
n	0	0	1	0	1	1	0	0

Popolazione iniziale con codifica binaria

1	5	1	4	8	2	6	3	7
2	3	6	2	1	7	4	5	8
3	8	3	5	1	2	6	4	7
4	2	4	8	3	1	7	6	5
5	3	7	1	6	5	8	2	4
...	7	5	3	8	6	4	1	2
n-1	8	6	4	5	2	7	3	1
n	5	7	8	4	1	2	6	3

Popolazione iniziale con codifica a permutazione

I passi della procedura

Inizializzazione e valutazione della fitness

Per problemi a massimizzare la fitness associata alla generica soluzione i , in generale, coincide con il suo valore di funzione obiettivo f_i . I valori vengono poi normalizzati in modo da associare a ciascuna soluzione una probabilità di selezione proporzionale al valore della fitness

Esempio calcolo fitness per problema a massimizzare

									f_i	$f_i / \sum f_i$
1	1	0	0	1	1	0	0	1	120	0.10
2	0	1	0	0	1	0	1	0	108	0.09
3	0	1	0	0	0	1	1	1	132	0.11
4	1	0	1	1	0	1	0	1	120	0.10
5	1	0	0	1	1	0	1	0	60	0.05
6	1	0	1	0	0	1	1	0	240	0.20
7	0	0	0	1	0	0	1	0	240	0.20
8	0	0	1	0	1	1	0	0	180	0.15
									Σf_i	1200

I passi della procedura

2 - Inizializzazione e valutazione della fitness

Per problemi a minimizzare la procedura va modificata fissando un valore $f_{\max} > \max f_i$ e ponendo la fitness associata alla generica soluzione i pari a $f'_i = f_{\max} - f_i$.

Esempio calcolo fitness per problema a minimizzare

Si pone $f_{\max} = 250 > 240$

									f_i	f'_i	$f'_i / \Sigma f'_i$
1	1	0	0	1	1	0	0	1	120	130	0.16
2	0	1	0	0	1	0	1	0	108	142	0.18
3	0	1	0	0	0	1	1	1	132	118	0.15
4	1	0	1	1	0	1	0	1	120	130	0.16
5	1	0	0	1	1	0	1	0	60	190	0.24
6	1	0	1	0	0	1	1	0	240	10	0.01
7	0	0	0	1	0	0	1	0	240	10	0.01
8	0	0	1	0	1	1	0	0	180	70	0.09
									Σf_i	1200	800

I passi della procedura

3 - Selezione

Consiste nella scelta di due elementi della popolazione.

Se $f_i/\Sigma f_i$ è il valore di fitness normalizzato associato all'elemento i , il metodo della roulette viene realizzato nel seguente modo:

- Si calcola per ciascun elemento i il valore cumulato di fitness normalizzato

$$F_i = \frac{\sum_{k=1,i} f_k}{\sum_{k=1,n} f_k}$$





- Si generano due numeri casuali a e b tra 0 e 1 ($0 \leq a \leq 1$; $0 \leq b \leq 1$);
- Si selezionano gli elementi s e t della popolazione tali che

$$F_{s-1} \leq a < F_s \qquad F_{t-1} \leq b < F_t$$

I passi della procedura

3 - Selezione

Ad esempio, con riferimento alla popolazione indicata, se si generano due numeri casuali sono $a=0.37$ e $b=0.68$, si selezionano le soluzioni 4 e 7

									f_i	$f_i / \Sigma f_i$	F_i		
	1	1	0	0	1	1	0	0	1	120	0.10	0.10	
	2	0	1	0	0	1	0	1	0	108	0.09	0.19	
	3	0	1	0	0	0	1	1	1	132	0.11	0.30	
	4	1	0	1	1	0	1	0	1	120	0.10	0.40	 $a=0.37$
	5	1	0	0	1	1	0	1	0	60	0.05	0.45	
	6	1	0	1	0	0	1	1	0	240	0.20	0.65	
	7	0	0	0	1	0	0	1	0	240	0.20	0.85	 $b=0.68$
	8	0	0	1	0	1	1	0	0	180	0.15	1.00	
									Σf_i	1200			

I passi della procedura

4- Generazione nuove soluzioni

Alla coppia selezionata si applicano gli **operatori genetici** allo scopo di generare nuove soluzioni del problema.

Gli operatori più frequentemente utilizzati sono il **crossover** e la **mutazione**.

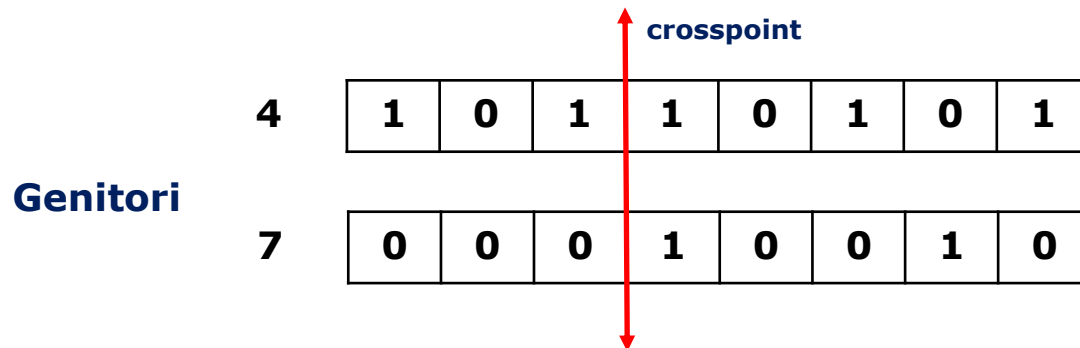
L'operatore di **crossover** accoppia due soluzioni (genitori) generandone altre due (figli) che presentano un patrimonio genetico e, quindi, valori delle variabili dedotti da quelli dei genitori.

La forma di implementazione del *crossover* dipende dalla codifica adottata per rappresentare le soluzioni

I passi della procedura

4 - Generazione nuove soluzioni

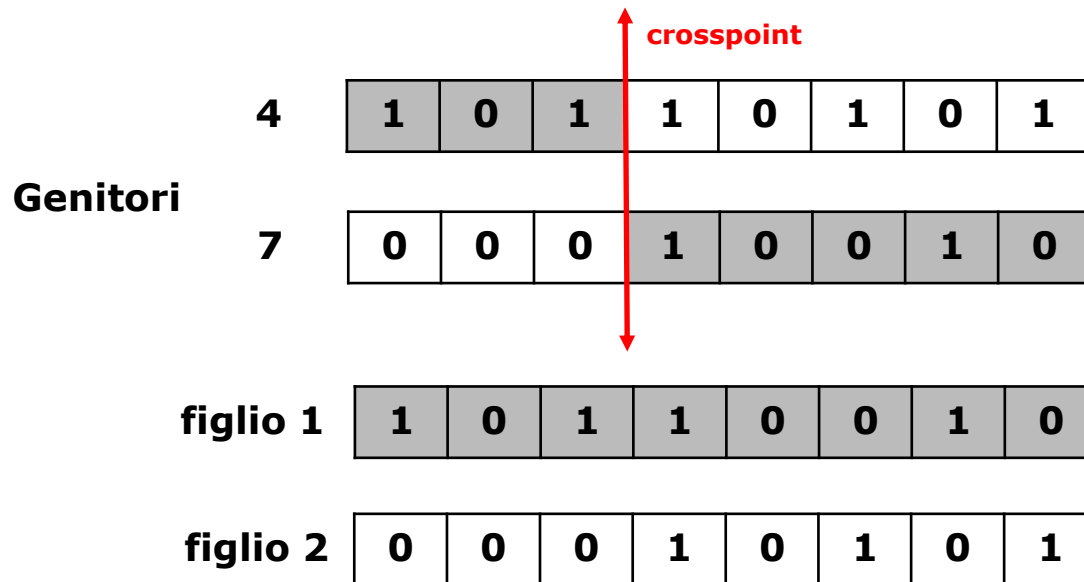
Nel caso di **codifica binaria**, si sceglie casualmente un punto di taglio (**crosspoint**), all'interno della coppia di stringhe rappresentative dei genitori, in modo da individuare due sottostringhe. I figli sono generati unendo la prima parte della prima sottostringa con la seconda parte della seconda sottostringa e viceversa.



I passi della procedura

4 - Generazione nuove soluzioni

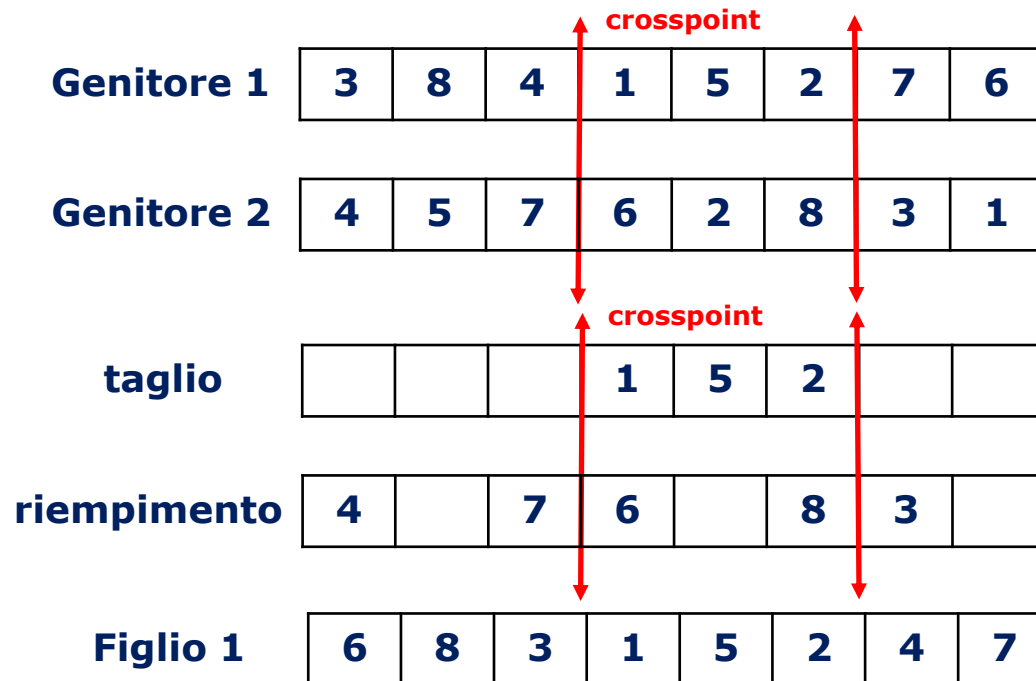
Nel caso di **codifica binaria**, si sceglie casualmente un punto di taglio (**crosspoint**), all'interno della coppia di stringhe rappresentative dei genitori, in modo da individuare due sottostringhe. I figli sono generati unendo la prima parte della prima sottostringa con la seconda parte della seconda sottostringa e viceversa.



I passi della procedura

4 - Generazione nuove soluzioni

Nella **codifica a permutazione**, attraverso due **crosspoint**, scelti casualmente, si individuano una sottostringa di "taglio" interna ai crosspoint in un genitore, e una sottostringa di "riempimento" ottenuta eliminando gli alleli della sezione di taglio, nell'altro. Un primo figlio viene generato riproducendo, nella stessa posizione, la sezione di taglio e aggiungendo gli alleli della sottostringa di riempimento, nell'ordine, a partire dal secondo crosspoints. Con lo stesso procedimento, invertendo il ruolo dei due genitori, si ottiene un'altra soluzione.



I passi della procedura

4 - Generazione nuove soluzioni

L'operatore di **mutazione** modifica casualmente il valore di un allele all'interno di una stringa.

Nel caso di codifica **binaria** si trasforma 0 in 1 o viceversa.

Nel caso di codifica a **permutazione** si scambia la posizione di due alleli

Fissata una probabilità di mutazione P_m di ciascun gene, si genera un numero casuale x ($0 \leq x \leq 1$): se x è $\leq P_m$ il gene viene mutato altrimenti resta inalterato. La probabilità di mutazione è, in generale, bassa ($0,01 \leq P_m \leq 0,02$).

Figlio

1	0	1	1	0	1	0	1
----------	----------	----------	----------	----------	----------	----------	----------

Numeri casuali

.63	.82	.18	.29	.71	.89	.09	.27
------------	------------	------------	------------	------------	------------	------------	------------

Esempio di mutazione

Con $P_m = 0,20$

I passi della procedura

4 - Generazione nuove soluzioni

L'operatore di **mutazione** modifica casualmente il valore di un allele all'interno di una stringa.

Nel caso di codifica **binaria** si trasforma 0 in 1 o viceversa.

Nel caso di codifica a **permutazione** si scambia la posizione di due alleli

Fissata una probabilità di mutazione P_m di ciascun gene, si genera un numero casuale x ($0 \leq x \leq 1$): se x è $\leq P_m$ il gene viene mutato altrimenti resta inalterato. La probabilità di mutazione è, in generale, bassa ($0,01 \leq P_m \leq 0,02$).

Figlio

1	0	1	1	0	1	0	1
----------	----------	----------	----------	----------	----------	----------	----------

Numeri casuali

.63	.82	.18	.29	.71	.89	.09	.27
------------	------------	------------	------------	------------	------------	------------	------------

Esempio di mutazione

Con $P_m = 0,20$

I passi della procedura

4 - Generazione nuove soluzioni

L'operatore di **mutazione** modifica casualmente il valore di un allele all'interno di una stringa.

Nel caso di codifica **binaria** si trasforma 0 in 1 o viceversa.

Nel caso di codifica a **permutazione** si scambia la posizione di due alleli

Fissata una probabilità di mutazione P_m di ciascun gene, si genera un numero casuale x ($0 \leq x \leq 1$): se x è $\leq P_m$ il gene viene mutato altrimenti resta inalterato. La probabilità di mutazione è, in generale, bassa ($0,01 \leq P_m \leq 0,02$).

Figlio

1	0	1	1	0	1	0	1
---	---	---	---	---	---	---	---

Numeri casuali

.63	.82	.18	.29	.71	.89	.09	.27
-----	-----	-----	-----	-----	-----	-----	-----

Esempio di mutazione

Con $P_m = 0,20$

Figlio

1	0	0	1	0	1	1	1
---	---	---	---	---	---	---	---

I passi della procedura

4 - Generazione nuove soluzioni

L'applicazione degli operatori genetici può produrre soluzioni non ammissibili. In tal caso, si può procedere in diversi modi.

- Si possono modificare gli operatori genetici al fine di evitare la produzione di soluzioni non ammissibili
- Le soluzioni non ammissibili possono essere sottoposte a meccanismi di "mutazione forzata" in modo da rientrare nell'ammissibilità
- Si assegnano alle soluzioni non ammissibili delle penalità nella fitness in modo che, introdotte comunque nella popolazione, presentino bassa probabilità di selezione ed alta probabilità di essere sostituite

I passi della procedura

5 - Sostituzione di elementi della popolazione

La popolazione evolve attraverso il progressivo rinnovo dei suoi elementi, ovvero introducendo gli individui generati in sostituzione di individui già presenti (in generale si tende a mantenere inalterata la dimensione della popolazione).

Una popolazione di n elementi può essere **aggiornata in modo totale o parziale**. Nel primo caso tutti gli n elementi sono sostituiti da nuovi elementi. L'aggiornamento parziale consiste, invece, nella sostituzione di $n' < n$ elementi.

- Si può confermare la presenza di elementi con i migliori valori di fitness e provvedere alla sostituzione degli altri.
- In alternativa si possono sostituire individui già presenti, introducendo i figli al posto dei genitori o al posto degli elementi con i peggiori valori di fitness.

I passi della procedura

6 - Criterio di arresto

Un algoritmo genetico converge quando gli individui della popolazione divengono più o meno simili. Al verificarsi di questa condizione l'operatore di crossover in pratica non è in grado di produrre nuove soluzioni e l'algoritmo esplora un sottoinsieme limitato dello spazio delle soluzioni.

Possibili criteri di arresto della procedura:

- Fissare un numero prestabilito di iterazioni
- Numero iterazioni nel corso delle quali la soluzione migliore della popolazione è rimasta la stessa.

La soluzione all'interno della popolazione finale che presenta il miglior valore di funzione obiettivo è assunta come soluzione finale dell'algoritmo.

I passi della procedura

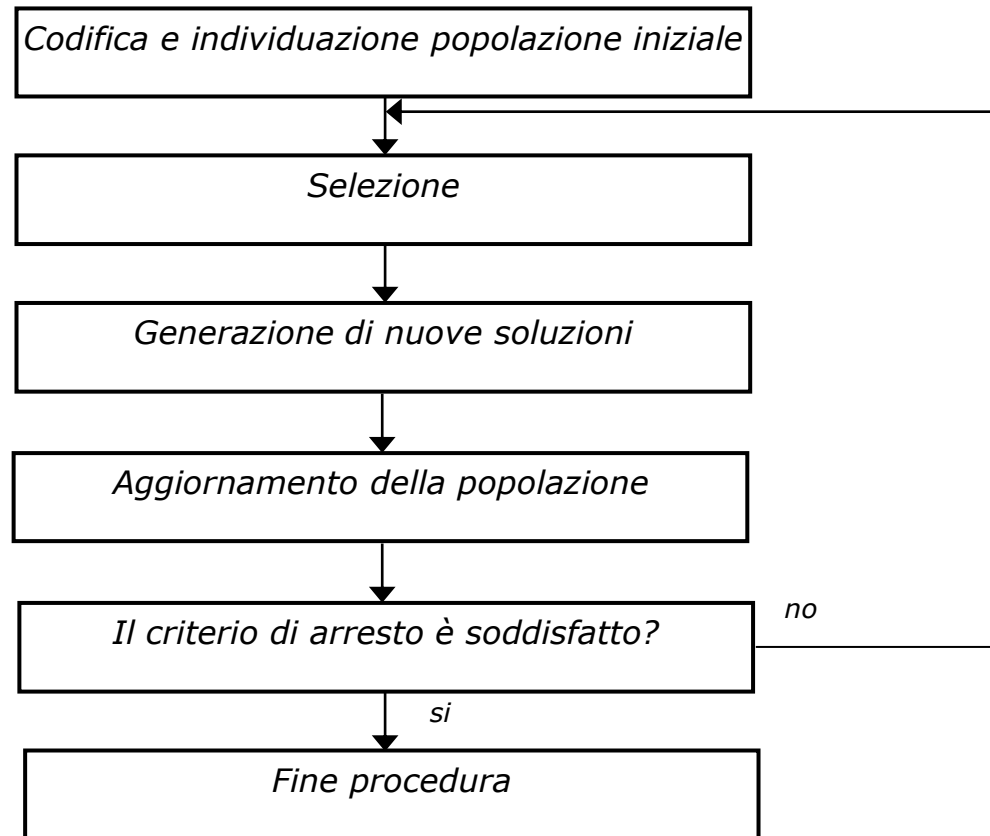
6 - Criterio di arresto

Nel corso dello sviluppo dell'algoritmo si può verificare un fenomeno di "**convergenza prematura**". In pratica, dopo un certo numero di iterazioni, la popolazione potrebbe risultare formata da individui molto simili con scarsa variabilità del patrimonio genetico. In questo caso l'algoritmo non riesce più ad individuare soluzioni migliori effettuando, così, iterazioni improduttive.

Per evitare questo fenomeno si possono introdurre forzatamente meccanismi che assicurino un certo grado di varietà nella popolazione alterando, ad esempio, il patrimonio genetico relativo ad un sottoinsieme della popolazione o modificando l'espressione della fitness in modo da accentuare diversità esistenti tra elementi della popolazione.

I passi della procedura

Schema logico di funzionamento della procedura



I passi della procedura

Caratteristiche generali della procedura

- ✓ In alcune applicazioni produce soluzioni molto buone
- ✓ Presenta una elevata flessibilità intesa come capacità di adattamento per la risoluzione di problemi diversi
- ✓ L'implementazione è molto semplice
- ✓ Non risulta molto sensibile alla scelta dei parametri di calibrazione
- ✓ I tempi di calcolo dipendono dal criterio di arresto
- ✓ La convergenza prematura avviene molto frequentemente
- ✓ Risultano più efficaci nella soluzione di problemi poco vincolati