CALCOLO EVOLUZIONISTICO ED

ALGORITMI GENETICI

Evoluzione

In ogni popolazione si verificano delle *mutazioni.*

In un ambiente che varia, le mutazioni possono generare individui che meglio si adattano alle nuove condizioni.

Questi individui tenderanno a sopravvivere più a lungo (*selezione naturale*) e quindi a generare un numero più elevato di figli (*riproduzione*).

I figli preservano in parte i caratteri genetici dei genitori (*cromosomi*, composti da *geni*) e, in parte, mescolandoli (*crossover*), creano nuovi tipi.

Il codice genetico che determina le caratteristiche di un individuo è detto *genotipo.* La manifestazione dei caratteri codificati in tale codice è detto *fenotipo*.

In questo modo, si ha una maggiore probabilità che nelle generazioni successive siano presenti i caratteri degli individui più “adatti” (con *fitness* più elevata) all’ambiente in cui vivono (*evoluzione*).

Il calcolo evoluzionistico:

Nelle scienze:

- Verifica tramite simulazioni di ipotesi in biologia, sociologia, religione ecc.

Per l'ingegneria:

- Ottimizzazione di funzioni
- Ottimizzazione combinatoria
- Apprendimento
- Soluzione di problemi

Corrispondenze natura-calcolo

Individuo	Soluzione di un problema
Popolazione	Insieme di soluzioni
Fitness	Qualità di una soluzione
Cromosoma	Rappresentazione di una soluzione
Gene	Componente di una rappresentazione
Crossover Mutazione	Operatori per la ricerca di soluzioni
Selezione Naturale	Riutilizzo di buone soluzioni

Calcolo evoluzionistico

Ispirandosi alla evoluzione naturale ed ai meccanismi esaminati, utilizzando le corrispondenze natura-calcolo descritte, si sviluppano metodi di ottimizzazione (ricerca della migliore soluzione possibile).

Generico algoritmo evoluzionistico

1. Inizializza una popolazione
2. Valuta la popolazione
3. Ripeti:
    a. seleziona un sottoinsieme della popolazione (*selezione*)
    b. Ricombina i geni degli individui selezionati (*Ricombinazione o crossover*)
    c. Modifica in modo casuale la popolazione ottenuta (*Mutazione*)
    d. Valuta la *fitness* della nuova popolazione

Algoritmi genetici

Negli algoritmi genetici si ottengono nuove soluzioni operando su una loro codifica: in termini genetici, si opera (come del resto accade anche in natura) sul solo *genotipo.*

E' quindi necessaria una decodifica da *genotipo* a *fenotipo*.

I *cromosomi* sono stringhe di simboli (es. 0 o 1).

Gli individui possono essere qualunque cosa possa essere codificata in una stringa di simboli.

I fenotipi possono essere, ad esempio, vettori di parametri, liste di possibili scelte, connessioni fra neuroni in una rete neurale ecc.

Semplice algoritmo genetico

1. Genera una popolazione random di cromosomi.

2. Decodifica ogni cromosoma per ottenere un individuo.

3. Valuta la fitness di ciascun individuo.

4. Genera una nuova popolazione, in parte *clonando* (copiando), in parte *ricombinando* e in parte *mutando* i cromosomi degli individui con la *fitness* più elevata.

5. Ripeti 2,3,4 fino ad una condizione di stop.

Problema della rappresentazione

Numeri

Interi (da 0 a 2^n-1, da K a K+N,
da 0 a M con $M \neq 2^n$-1)

Reali

Elementi appartenenti ad insiemi limitati

Vettori di parametri o numeri

Problema della codifica

Codifica di Gray

I numeri interi da 0 a 2^n-1 sono codificati con una stringa binaria di lunghezza n.
Interi consecutivi hanno una rappresentazione che differisce per un solo bit

Es.

0	000
1	001
2	011
3	010
4	110
5	111
6	101
7	100

Vantaggio
Una qualsiasi inversione di un bit produce piccoli cambiamenti

Svantaggio
Quando il cambiamento è grande è maggiore che con una normale codifica binaria

Selezione

E' l'operazione mediante la quale gli individui (in realtà il loro genotipo, rappresentato dai cromosomi) sono selezionati per la riproduzione.

Per simulare la selezione naturale gli individui con fitness più elevata hanno maggiore probabilità di essere selezionati.

Vi sono diverse strategie per la selezione (alcune non plausibili biologicamente).

Di solito in un algoritmo genetico:

a. Un gruppo di soluzioni è selezionato per l'accoppiamento (*mating pool*)
b. Coppie di individui sono estratti a caso dal mating pool e vengono accoppiati

Ipotesi fondamentali per la selezione

1. Esiste una misura Q della bontà di una soluzione.

2. E’ sempre positiva

3. Va massimizzata

4. La Q di un individuo è la sua *fitness*

Selezione proporzionale alla fitness (*fitness-proportionate selection*)

E' il metodo di selezione più comune.

Ad ogni individuo è assegnata una probabilità di essere selezionato

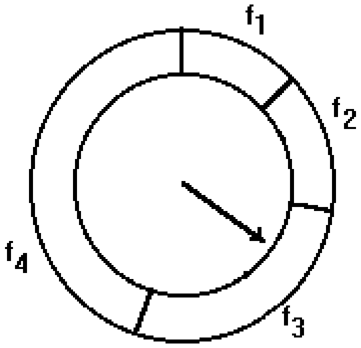
$$p_i = f_i/\Sigma_j f_j$$

Implementazione

Supponiamo di avere 4 individui con fitness

$f_1=f_2=10 \qquad f_3=15 \qquad f_4=25$

1. Metodo della roulette



Ogni posizione della freccia corrisponde ad un certo numero. Si estrae un numero casuale e si seleziona l'individuo puntato dalla freccia.

2. Vettore di dimensione N

112233344444
0 N-1

Si estrae un numero a caso (da 0 a N-1) e si seleziona l'individuo corrispondente al contenuto del vettore nella posizione estratta.

3. Numero reale compreso fra 0 e $\Sigma_j f_j$

Si estrae un numero a caso r compreso in tale intervallo e si seleziona l'individuo i tale che

$$\Sigma_{j=1,i-1} f_j \leq r < \Sigma_{j=1,i} f_j$$

Problemi

Convergenza prematura

Se un individuo i ha fitness molto maggiore della media della popolazione ma molto minore della massima fitness possibile, tenderà a essere sempre selezionato e quindi a generare una popolazione “mediocre”.

Stagnazione

Dopo un certo numero di generazioni, tutti gli individui hanno una buona fitness e quindi tendono ad avere la stessa probabilità di essere selezionati.

Selezione per rango (*rank selection*)

Si ordinano gli individui in ordine decrescente di fitness e si attribuisce una probabilità decrescente (ad es. linearmente o esponenzialmente) in funzione della posizione in classifica, indipendentemente dal valore della fitness.

Vantaggi

Non si ha convergenza prematura (nessun individuo ha probabilità molto maggiore degli altri di essere selezionato)

Non c'è stagnazione: le probabilità sono sempre ugualmente distribuite.

Svantaggi

Computazionalmente pesante.

Nota: non è biologicamente plausibile.

Selezione tramite torneo (*tournament selection*)

Per ogni individuo da selezionare, si seleziona un gruppo di individui e si clona il migliore (torneo).

Vantaggi

Come selezione per rango, ma non c'è necessità di ordinamento.

Selezione elitista

Almeno una copia dell'individuo migliore viene mantenuta nella generazione successiva.

Vantaggi

Non si perdono soluzioni buone nei processi "casuali" di selezione.

Svantaggi

Si possono raggiungere massimi locali se i caratteri di tale individuo diventano "dominanti".

CROSSOVER SINGOLO-PUNTO

10001|10011101010 11100|01000111001 GENITORI

10001|01000111001 11100 |10011101010 FIGLI

CROSSOVER DOPPIO-PUNTO

10001|1001110|1010 11100|0100011|1001 GENITORI

11100|1001110|1001 10001 |0100011|1010 FIGLI

CROSSOVER UNIFORME

Ogni "gene" è selezionato a caso da uno dei due genitori

MUTAZIONE

1000110011101010 1000110010101010