

Intelligenza Artificiale

Algoritmi Genetici

Alberto Broggi

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*).

Alberto Broggi

Algoritmi Genetici

Evoluzione

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

Alberto Broggi

Algoritmi Genetici

Evoluzione: nomenclatura

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

Alberto Broggi

Algoritmi Genetici

Il calcolo evoluzionistico

- Nelle scienze:
 - Utilizzo di simulazioni per la verifica di ipotesi in biologia (ad es. nascita e sviluppo di organismi viventi), sociologia, ecc.
- Nell'ingegneria:
 - Ottimizzazione di funzioni
 - Ottimizzazione combinatoria
 - Apprendimento
 - Soluzione di problemi

Alberto Broggi

Algoritmi Genetici

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 |

Alberto Broggi

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).
- Le prime idee risalgono agli anni '60, ma hanno trovato applicazione solo negli anni '80

Alberto Broggi

Generico algoritmo evoluzionistico

1. Inizializza una popolazione
2. 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

Alberto Broggi

Algoritmi genetici

- Negli algoritmi genetici si ottengono nuove soluzioni operando su una loro codifica: in termini genetici, si opera sul solo *genotipo* (come del resto accade anche in natura).
- E' quindi necessaria una decodifica da *genotipo* a *fenotipo*.

Alberto Broggi

Algoritmi genetici

- I *cromosomi* sono stringhe di simboli (es. 0 o 1).
- Gli *individui* possono essere qualunque cosa che possa essere codificata con 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.

Alberto Broggi

Semplice algoritmo genetico

1. Genera una popolazione casuale 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.

Alberto Broggi

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

Alberto Broggi

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
- Vantaggio:
 - Una qualsiasi inversione di un bit produce piccoli cambiamenti
- Svantaggio:
 - Quando il cambiamento è grande è maggiore che con una normale codifica binaria

Es.

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

Alberto Broggi

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 maggior probabilità di essere selezionati.

Alberto Broggi

Selezione

- Vi sono diverse strategie per la selezione (tra cui 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

Alberto Broggi

Ipotesi fondamentali per la selezione

- Esiste una misura Q della bontà di una soluzione.
- E' sempre positiva
- Va massimizzata
- La Q di un individuo è definita come la sua *fitness*

Alberto Broggi

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

- E' il metodo di selezione più comune.
- Ad ogni individuo è assegnata una probabilità di essere selezionato

$$p_i = \frac{f_i}{\sum f_i}$$

Implementazione:

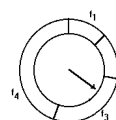
Supponiamo di avere 4 individui con fitness

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

Alberto Broggi

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.



Alberto Broggi

Vettore di dimensione N

1 1 2 2 3 3 3 4 4 4 4
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.

Alberto Broggi

Numero reale compreso fra 0 e $\sum f_j$

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

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

Alberto Broggi

Problemi

- *Convergenza prematura*
 - Se un individuo i ha fitness molto maggiore della media della popolazione ma molto minore della massima fitness possibile, tenderà ad 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.

Alberto Broggi

Selezione per rango (rank selection)

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

Alberto Broggi

Selezione per rango (rank selection)

- *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.

Alberto Broggi

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.

Alberto Broggi

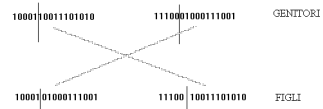
Selezione elitista

- 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".

Alberto Broggi

Esempi

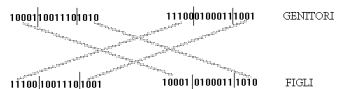
CROSSOVER SINGOLO-PUNTO



Alberto Broggi

Esempi

CROSSOVER DOPPIO-PUNTO



Alberto Broggi

Esempi

CROSSOVER UNIFORME

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

MUTAZIONE

100011001101010 100011001101010

Alberto Broggi

Condizione di stop

- Il processo viene arrestato:
 - quando si raggiungono soluzioni con fitness sufficientemente elevata
 - quando la fitness media cresce al di sotto di una soglia (quando, cioè, tutti gli individui sono simili tra loro)
 - dopo un numero definito di iterazioni
- Il risultato può essere:
 - il migliore dell'ultima generazione
 - il migliore tra tutti gli individui generati dall'inizio del processo (situazione non plausibile geneticamente)

Alberto Broggi

Pacchetti software

- Nei pacchetti software si ha la possibilità di definire
 - il *genotipo* (la codifica)
 - i meccanismi evolutivi
- Rimane a carico dell'utente:
 - decodificare il *genotipo* per ottenere il *fenotipo*
 - calcolare la *fitness* del *fenotipo*
 - definire i parametri che regolano l'evoluzione
 - » dimensione della popolazione
 - » frequenza del crossover
 - » frequenza della mutazione

Alberto Broggi