

Alessandro Bollini
bollini@vision.unipv.it

Dipartimento di Informatica e Sistemistica
Università di Pavia
Via Ferrata, 1
27100 Pavia

Algoritmi Genetici



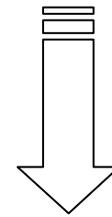
Algoritmo genetico

- Algoritmo evolutivo.
- Modello evolutivo a livello di genotipo del singolo individuo.
- Genotipo generalmente costituito da una stringa binaria di lunghezza fissa.
- Operatori di riproduzione, ricombinazione e mutazione.

Rappresentazione

- Genotipo: generalmente una stringa binaria di lunghezza fissa.
- Fenotipo: insieme di parametri di controllo per la soluzione.
- Mappatura genotipo-fenotipo.

0 0 1 0 0 1 1



$X[0]=0.9$

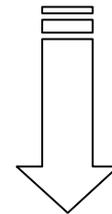
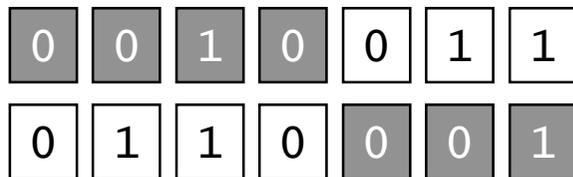
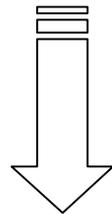
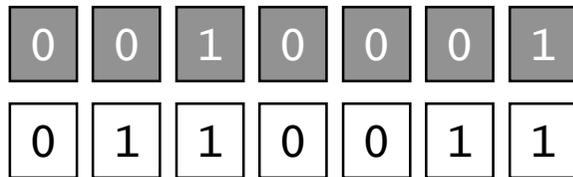
$x[1]=1.5$

...

$X[N]=0.3$

Operatori genetici

- Riproduzione.
- Ricombinazione (*crossover*) e mutazione.



Configurazione

- Obiettivo.
- Variabili di controllo.
- Dimensioni del genotipo.
- Mappatura genotipo-fenotipo.
- Funzione di valutazione delle prestazioni.
- Criteri di terminazione.
- Dimensione della popolazione.
- Probabilità di applicazione degli operatori.

Algoritmo di base

- Creazione della popolazione iniziale.
- Fino al raggiungimento dei criteri di terminazione:
 - Crea una nuova popolazione:
 - Seleziona uno o più individui dalla popolazione corrente.
 - Crea un nuovo individuo applicando gli operatori genetici.
 - Valuta le prestazioni del nuovo individuo.
 - Inserisci il nuovo individuo nella nuova popolazione.

Dinamica della popolazione

- La popolazione rappresenta un campione dello spazio delle soluzioni.
- Ogni iterazione dell'algoritmo crea un nuovo campione partendo dal vecchio, tramite l'applicazione degli operatori selettivi e genetici.
- Empiricamente si osserva che la *fitness* media della popolazione cresce, cioè che il campione si addensa in aree a *fitness* alta.

Allocazione ottima delle prove

- Ipercubi nello spazio delle soluzioni descritto dalla stringa binaria.
- Stima della *fitness* media sull'ipercubo.
- Riallocazione del campione proporzionalmente allo scostamento dalla *fitness* media sull'intero spazio:
$$d(k, t+1) = d(k, t) * f(k) / f$$
- Come stimare la media sugli ipercubi?
- Come generare il nuovo campione?

Parallelismo implicito

- Ogni stringa binaria di N bit appartiene a 2^N tra 3^N possibili ipercubi.
- Una popolazione di I individui di N bits fornisce $I \cdot 2^N$ dati per la stima della *fitness* negli ipercubi.

Selezione proporzionale

- Selezione proporzionale alla *fitness*:
$$p(i, t+1) = p(i, t) * f(i) / f$$
- La riproduzione semplice con selezione proporzionale alla *fitness* genera nuovi campioni in cui la distribuzione degli individui negli ipercubi è ottimale secondo quanto definito.
- La riproduzione semplice non genera varietà e quindi non è in grado di esplorare nuove soluzioni.
- La finitezza del campione genera errori di campionamento (*genetic drift*).

L'effetto degli operatori

- Gli operatori genetici ripristinano la varietà del campione a costo di perturbare il campionamento degli ipercubi: alcuni dei blocchi selezionati possono essere distrutti e quindi non propagati.
- L'effetto è ridotto nel caso di *blocchi costruttivi* compatti.
- Operatori di inversione e di *linkage* possono essere introdotti per promuovere la compattezza dei blocchi.

Convergenza e prestazioni

- La riproduzione e la selezione proporzionale alla *fitness* implementano implicitamente la strategia ottima per lo *sfruttamento* delle soluzioni già esplorate.
- Gli operatori genetici forniscono la varietà richiesta per l'*esplorazione* di nuove soluzioni.
- Il peso relativo deve essere calibrato manualmente.

Non-linearità

- La *fitness* è stimata indipendentemente su ogni ipercubo.
- I blocchi costruttivi che sfruttano le non-linearità per ottenere *fitness* alte verranno progressivamente sfruttati.
- La stima della fitness diventa progressivamente più precisa, a condizione che la popolazione sia sufficientemente ampia.
- Eventuali errori stocastici nella stima possono essere corretti.

Conclusioni

- Il parallelismo implicito degli algoritmi genetici permette di esplorare vaste aree dello spazio di ricerca con relativamente pochi individui.
- La selezione proporzionale alla *fitness* implementa la strategia ottima per lo sfruttamento dello spazio esplorato.
- Il parallelismo implicito permette di gestire le interazioni non-lineari.