CS 4752
Introduction to Computational Intelligence
Lecture 6
September 20, 2011

Outline

• Population Model
  – Generational: non-overlapping population
  – Steady-state: overlapping population
• Selection Methods
  – Stochastic Selection
  – Deterministic Selection
Evolutionary Algorithm Workflow

When to Select

• Selection can occur at two stages of evolution:
  – Parent selection: select individuals from the current population to take part in mating.
  – Survivor selection: select individuals from offspring and/or current population to generate new population.

• Selection operators work on the entire individual, i.e. they are representation-independent.
Generational Model

• Non-overlapping Population:

\[ \text{gen-0} \rightarrow \text{gen-1} \rightarrow \text{gen-2} \rightarrow \cdots \rightarrow \text{gen-n} \]

• Parent selection only:
  • All offspring are kept in the following generation – Canonical Genetic Algorithms
• Offspring selection with random parent selection:
  • Random parent selection to generate a large number of offspring;
  • Fitter offspring are selected to form the new generation - Evolution Strategies \((\mu, \lambda)\).

Generational Model - Continue

• Under stochastic selection, an individual, regardless how fit it is, may only live for one generation, hence has a short-term impact on evolution.
• During stochastic selection, best solutions might get lost and are never carried over to new generation.
• These can be fixed by deterministic selection, or…
Steady-state Model

- Overlapping Population

\[ \text{pop} \rightarrow \text{parents} \]
\[ \text{replace} \rightarrow \text{offspring} \]

- Generation Gap:
  - The proportion of the population that is replaced [Sarma & De Jong, 1995].
  - Generational model: \( \text{pop\_size}/\text{pop\_size}=1.0 \)
  - Steady-state model: \( \#_{\text{replaced}}/\text{pop\_size} \)

Steady-state Model - Continue

- Offspring and parents compete to survive in the population.
- Fit individual can live for a long period of time to impact the evolution.
- A fit offspring can have impact on the evolution immediately after its birth, without waiting until the next generation.
- Advantageous or disadvantageous for evolutionary search?
How to Select

- Stochastic Selection (probabilistic)
  - Random (stochastic uniform)
  - Fitness proportionate
  - Rank (linear vs. non-linear)
  - Tournament (binary vs. more)
- Deterministic Selection
  - Deterministic uniform
  - Deterministic replacement
  - Truncation (the best n)

Selection Pressure

- Selection methods are characterized by their selection pressure, also referred to as the takeover time, which relates to the time it requires to produce a uniform population.
- This can be estimated by repeated application of the selection method alone (w/o genetic operation).
- Higher selection pressure: population diversity (number of unique individual) is decreased more rapidly.
Random Selection

• Non-fitness based selection
• Each individual in the population has the same probability to be selected (stochastic random).
• Random selection has the lowest selection pressure among all selection methods.
• This can be used to pair with a non-random selection for parent/offspring selection.

Fitness Proportionate (Roulette Wheel)

• The probability of an individual $i$ to be selected from a population of $n$ individuals is

$$P_i = \frac{f_i}{\sum_{i=1}^{n} f_i}$$

• The expected number of each individuals to be selected is

$$E_i = \frac{f_i}{\bar{f}}$$

– $f_i$: non-negative fitness of individual $i$
– $\bar{f}$: average population fitness
– maximization problem (*-1 to convert minimization to maximization)
Example

<table>
<thead>
<tr>
<th>id</th>
<th>f(x)</th>
<th>$P_i$</th>
<th>$E_i$</th>
<th>Actual Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>169</td>
<td>0.09204793</td>
<td>0.368191721</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1276</td>
<td>0.694989107</td>
<td>2.779956427</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>0.016339869</td>
<td>0.065359477</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>361</td>
<td>0.196623094</td>
<td>0.786492375</td>
<td>1</td>
</tr>
<tr>
<td>sum</td>
<td>1836</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>average</td>
<td>459</td>
<td>0.25</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>max</td>
<td>1276</td>
<td>0.694989107</td>
<td>2.779956427</td>
<td>3</td>
</tr>
</tbody>
</table>

over select the super individual 2.

Shortcoming

- During early stage of evolution, super individuals may take over a significant proportion of the population and cause premature convergence.
- Later state of evolution, the fitness of all individuals are close to each other, which makes the fitness-based selection method act like random selection (weak selection pressure)
Fitness Scaling

- Linear scaling: \( f'(x) = af(x) + b \)
  - The average population members contribute one expected count to the next generation.
  - The best population members contributes \( c \) expected counts to the next generation, \( 1.2 \leq c \leq 2 \).

\[
a = \frac{(c - 1) \cdot \text{avg}(f)}{\text{max}(f) - \text{avg}(f)}
\]

\[
b = \frac{\text{avg}(f) \cdot (\text{max}(f) - c \cdot \text{avg}(f))}{\text{max}(f) - \text{avg}(f)}
\]

### Linear Scaling

<table>
<thead>
<tr>
<th>id</th>
<th>f(x)</th>
<th>a*f(x)+b</th>
<th>pi</th>
<th>Ei</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>169</td>
<td>296.07</td>
<td>0.16</td>
<td>0.645</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1276</td>
<td>918</td>
<td>0.5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>217.98</td>
<td>0.1187</td>
<td>0.4749</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>361</td>
<td>403.94</td>
<td>0.22</td>
<td>0.88</td>
<td>1</td>
</tr>
<tr>
<td>sum</td>
<td>1836</td>
<td>1836</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>average</td>
<td>459</td>
<td>459</td>
<td>0.25</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>max</td>
<td>1276</td>
<td>918</td>
<td>0.5</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

\( c=2; \)
Rank Selection

- Selection probabilities are based on relative rather than absolute fitness.
- Individuals are sorted according to their fitness.
- The probability of an individual to be selected is proportionate to its rank.
- Ranking introduces a uniform scaling across the population and provides a simple and effective way of controlling selective pressure.

Linear Rank Selection

- For population size n, rank 1 has the best fitness and rank n has the worst fitness:

  \[ p_1 = \frac{2}{n} - \varepsilon \]
  \[ p_n = \varepsilon \]
  \[ p_i = \left( \frac{2}{n} - \varepsilon \right) - \left( \frac{2}{n} - 2\varepsilon \right) \frac{i - 1}{n - 1} \]

\( \varepsilon \) controls the slope of the linear probability distribution by ranging from 0.0 (max slope) to 1/n (0 slope, flat line).
Linear Ranking Selection

<table>
<thead>
<tr>
<th>ind</th>
<th>fitness</th>
<th>rank</th>
<th>$P_i$</th>
<th>$E_i$</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>169</td>
<td>3</td>
<td>0.1667</td>
<td>0.6668</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1276</td>
<td>1</td>
<td>0.4999</td>
<td>1.9996</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>4</td>
<td>1E-04</td>
<td>0.0004</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>361</td>
<td>2</td>
<td>0.3333</td>
<td>1.3332</td>
<td>1</td>
</tr>
<tr>
<td>sum</td>
<td>1836</td>
<td>10</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>average</td>
<td>459</td>
<td>2.5</td>
<td>0.25</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>max</td>
<td>1276</td>
<td>4</td>
<td>0.4999</td>
<td>1.9996</td>
<td>2</td>
</tr>
</tbody>
</table>

$\varepsilon = 0.001$

Comparison

- fitness-proportional
- linear-scaling
- Linear-Ranking
Tournament Selection

• A small subset of $n$ individuals are chosen at random (with replacement), and best individuals in this set is/are selected.
• The larger the tournament size $n$, the stronger the selection pressure.
• If pick without replacement (each individual can only be selected once the most), selection pressure is increased or decreased?

Binary-Tournament

• Two individuals are selected randomly and the one with better fitness is the winner.
• Easy to implement, as there is no need to maintain the order of the individuals in the population, which is required by other stochastic selection methods.