Nature Inspired Search and Optimisation Advanced Aspects of Nature Inspired Search and Optimisation

Lecture 6: Selection and reproduction

Shan He

School for Computational Science
University of Birmingham

January 30, 2020

Outline of Topics

1. Selection
2. Reproduction
3. Conclusion

Generic Evolutionary Algorithm

Generic Evolutionary Algorithm

$\mathbf{X}_0$:= generate initial population of solutions
terminationflag := false
t := 0
Evaluate the fitness of each individual in $\mathbf{X}_0$.
while (terminationflag != true)
    Selection: Select parents from $\mathbf{X}_t$ based on their fitness.
    Variation: Breed new individuals by applying variation operators to parents
    Fitness calculation: Evaluate the fitness of new individuals.
    Reproduction: Generate population $\mathbf{X}_{t+1}$ by replacing least-fit individuals
    t := t + 1
    If a termination criterion is met: terminationflag := true
Output x_{best}

Selection

- Selection is not a search operator but influences search performance significantly
- Selection usually is performed before variation operators: selects better fit individuals for breeding
- Selection: emphasises on **exploiting** better solutions in a population:
  - Select one or more copies of good solutions.
  - Inferior solutions will be selected but with a much less chance
- **Question**: Why we still select those inferior solutions?

Selection schemes

- Selection schemes:
  - Fitness Proportional Selection
  - Ranking Selection
  - Truncate selection
  - Tournament Selection
  - $(\mu + \lambda)$ and (μ, λ) selection
- Can be grouped based on
  - Relative fitness: fitness proportional selection
  - Ranking: ranking, tournament, truncation, $(\mu + \lambda)$ and (μ, λ)

Fitness Proportional Selection

- Fitness Proportional Selection = roulette wheel selection
- Selecting individual i with a probability:

$$p_i = \frac{f_i}{\sum_{j=1}^{M} f_j},$$

where f_i is the fitness value of individual i, M is the number of individual

- Does not allow negative fitness value
- Individual with higher fitness values will be more likely to be selected, but still a chance that they may be eliminated
- Individual with low fitness values may survive the selection process: allows some weak individuals who may help escaping from local optima

Fitness Proportional Selection: scaling

- Observation: in early generations, there might be a domination of “super individuals” with very high fitness values
- **Question**: what problems will “super individuals” cause in EAs?
- Observation: In later generations, there might be no much separation among individuals
- **Question**: What are the problems in EAs with no much separation among individuals?

Fitness Proportional Selection: scaling

- Problems:
  - **Question**: what problems will "super individuals" cause in EAs?
  - **Answer**: Premature convergence to a local optimum
  - **Question**: What problems will caused in EAs with no much separation among individuals?
  - **Answer**: Slow convergence
- Question: How to maintain the same selection pressure throughout the run
- Solution: replace raw fitness values f_i with a scaled fitness value f_i'
- **Linear scaling**:

$$f_i' = a + b \cdot f_i$$

where a and b are constants, usually set as $a = \max(\mathbf{f})$ and $b = \min(\mathbf{f})/M < 1$, where $\mathbf{f} = f_1, f_2, \cdots, M$

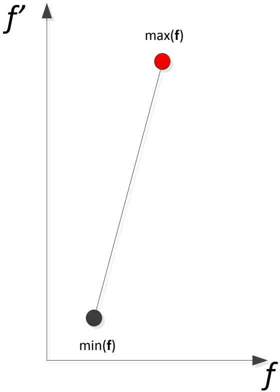


Figure: Fitness scaling using linear scaling

Ranking Selection

- Sort population size of M from best to worst according to their fitness values:

$$x'_{(M-1)}, x'_{(M-2)}, x'_{(M-3)}, \cdots, x'_{(0)},$$

- Select the top γ-ranked individual x'_γ with probability $p(\gamma)$, where γ is the rank and $p(\gamma)$ is a ranking function, e.g.
  - linear ranking
  - exponential ranking
  - power ranking
  - geometric ranking

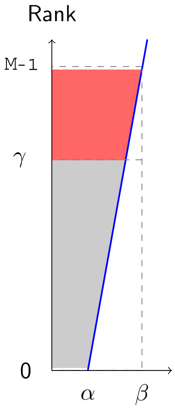
Linear ranking function

- **Linear ranking function:**

$$p(\gamma) = \frac{\alpha + (\beta - \alpha) \cdot \frac{\gamma}{M-1}}{M}$$

where $\sum_{\gamma=0}^{M-1} p(\gamma) = 1$ implies $\alpha + \beta = 2$ and $1 \leq \beta \leq 2$

- In expectation
  - best individual, i.e., $\gamma = M - 1$, reproduced β times: $p(M-1) = \frac{\beta}{M}$
  - worst individual, i.e., $\gamma = 0$, reproduced α times: $p(0) = \frac{\alpha}{M}$
- **Question**: how to set α and β to make EA behave like a random search algorithm?



Other ranking functions

Give you different, usually non-linear relationships between the rank γ and the selection probability $p(\gamma)$:

- **Power ranking function**:

$$p(\gamma) = \frac{\alpha + (\beta - \alpha) \cdot (\frac{\gamma}{M-1})^k}{C}$$

- **Geometric ranking function**:

$$p(\gamma) = \frac{\alpha \cdot (1 - \alpha)^{M-1-\gamma}}{C}$$

- **Exponential ranking function**:

$$p(\gamma) = \frac{1 - e^{-\gamma}}{C}$$

where C is a normalising factor and $0 < \alpha < \beta$.

Truncation Selection

- Steps:
  - Rank individuals by fitness values
  - Select some proportion, k, (e.g. $k = \frac{1}{2}$), of the top ranked individuals
- Usually, $k = 0.5$ (top 50%) or $k = 0.3$ (top 30%)
- Can be seen as the simplest, deterministic ranking selection

Tournament Selection

- Tournament selection with tournament size k:
  - Step 1: Randomly sample a subset P' of k individuals from population P
  - Step 2: Select the individual in P' with highest fitness
  - Repeat Steps 1 and 2 until enough offspring are created
- One of the most popular selection methods in genetic algorithms.
- Binary tournament selection $(k = 2)$ is the most popular one

$(\mu + \lambda)$ and (μ, λ) selection

- First proposed in Evolution Strategies.
- $(\mu + \lambda)$ selection:
  - Parent population of size μ
  - Generate λ offspring from randomly chosen parents
  - Select μ best individuals **among parents and offspring**
- (μ, λ) selection where $\lambda > \mu$
  - Parent population of size μ
  - Generate λ offspring from randomly chosen parents
  - Select μ best individuals **among offspring**

Selection pressure

- Selection pressure: degree to which selection emphasises on the better individuals.
- **Question 1:** How does selection pressure affect the balance between exploitation and exploration?
- **Question 2:** Given an EA with a selection scheme, how will you measure selection pressure?

Selection pressure

- **Question 1:** How does selection pressure affect the balance between exploitation and exploration?
- **Answer:**
  - Higher selection pressure $\rightarrow$ exploitation $\rightarrow$ fast convergence to local optimum, e.g., premature
  - Low selection pressure $\rightarrow$ exploration $\rightarrow$ slow convergence

Selection pressure

- **Question 2:** Given an EA with a selection scheme, how will you measure selection pressure?
- **Answer**: Take-over time τ^* [1]:
  - Let's assume population size is M and initial population with one unique fittest individual x^*
  - Apply selection repeatedly with no other operators.
  - τ^* is number of generations until population consists of x^* only.
- Higher take-over time means lower selection pressure.

[1] Goldberg, D. E. and Deb, K. (1991). A comparative analysis of selection schemes used in genetic algorithms. In Foundations of Genetic Algorithms. Morgan Kaufmann.

Selection pressure

Selection function	$\tau^* \approx$	Note
Fitness prop.	$\frac{M \ln M}{c}$	Assuming a power law objective function $f(x) = x^c$
Linear ranking	$\frac{2 \ln(M-1)}{\beta-1}$	$1 \leq \beta \leq 2$
Truncation	$\ln M$	
Tournament	$\frac{\ln M}{\ln k}$	tournament size k
$(\mu + \lambda)$	$\frac{\ln \lambda}{\ln(\lambda/\mu)}$	

Reproduction

- Reproduction: to control how genetic algorithm creates the next generation
- Generational vs Steady state
  - Generational EAs: also called standard EAs, use all new individuals after variations to replace the worse individuals in the old population to create a new population (then selection)
  - Steady state EAs: only use **a few or even one single new individual** to replace the old population at any one time
- n-Elitisms: always copy the n best individuals to the next generation
- Generational gap: the overlap (i.e., individuals that did not go through any variation operators) between the old and new generations.

Conclusion

- Selection and reproduction are two important ingredients of EAs
- The major difference between selection methods is based on:
  - Relative fitness: fitness proportional selection
  - Ranking: ranking, tournament, truncation, $(\mu + \lambda)$ and (μ, λ)
- Takeover time: quantitative measure of selection pressure
- Selection pressure is used to control the balance between exploitation and exploration:
  - Higher selection pressure $\rightarrow$ exploitation $\rightarrow$ fast convergence to local optimum, e.g., premature
  - Low selection pressure $\rightarrow$ exploration $\rightarrow$ slow convergence