Computational Intelligence

Unit # 4

Acknowledgement

- The slides of this lecture have been taken from the lecture slides of "CSE659 – Computational Intelligence" by Dr. Sajjad Haider.

Announcements

- Quiz 1 in the next class - this will cover the topics discussed so far in the class
- Assignment 1 will also be release by the end of this week

Selection

- Selection is one of the main operators in EAs, and relates directly to the Darwinian concept of survival of the fittest.

General Pattern

- A population of constant size μ is evolved over time.

- The current population is used as a source of parents to produce λ offspring.

- The expanded population is reduced from $\mu+\lambda$ to μ individuals.

Selection Procedure

- When selection is used to choose which individuals reproduce, the process is referred to as ***pre-selection*** (parent(s) selection).

- When it is used to select the individuals that survive to the next generation it is called ***post-selection*** (survival selection).

Parent Selection Mechanism

- Assigns variable probabilities of individuals acting as parents depending on their fitness
- Usually probabilistic
  - high quality solutions more likely to become parents than low quality
  - but not guaranteed
  - even worst in current population usually has non-zero probability of becoming a parent
- This *stochastic* nature can aid escape from local optima

Selection Schemes

- Fitness Proportional
- Rank Selection
- Tournament Selection
- Truncation
- Elitist
- Uniform Stochastic

Fitness Proportional Selection

- Proportional selection, proposed by Holland, biases selection towards the most fit individuals.

- A probability distribution proportional to the fitness is created, and individuals are selected by sampling the distribution.

- Because selection is directly proportional to fitness, it is possible that strong individuals may dominate in producing offspring, thereby limiting the diversity of the new population. This is known as **premature convergence**.

- In other words, proportional selection has a high selective pressure.

Fitness Proportional Selection (Cont'd)

- When fitness values are all very close together, there is almost no selection pressure. Therefore, later in a run, when some convergence has taken place and the worst individuals are gone, the performance only increases very slowly.

- Also known as Roulette Wheel Selection.

Roulette Wheel Selection
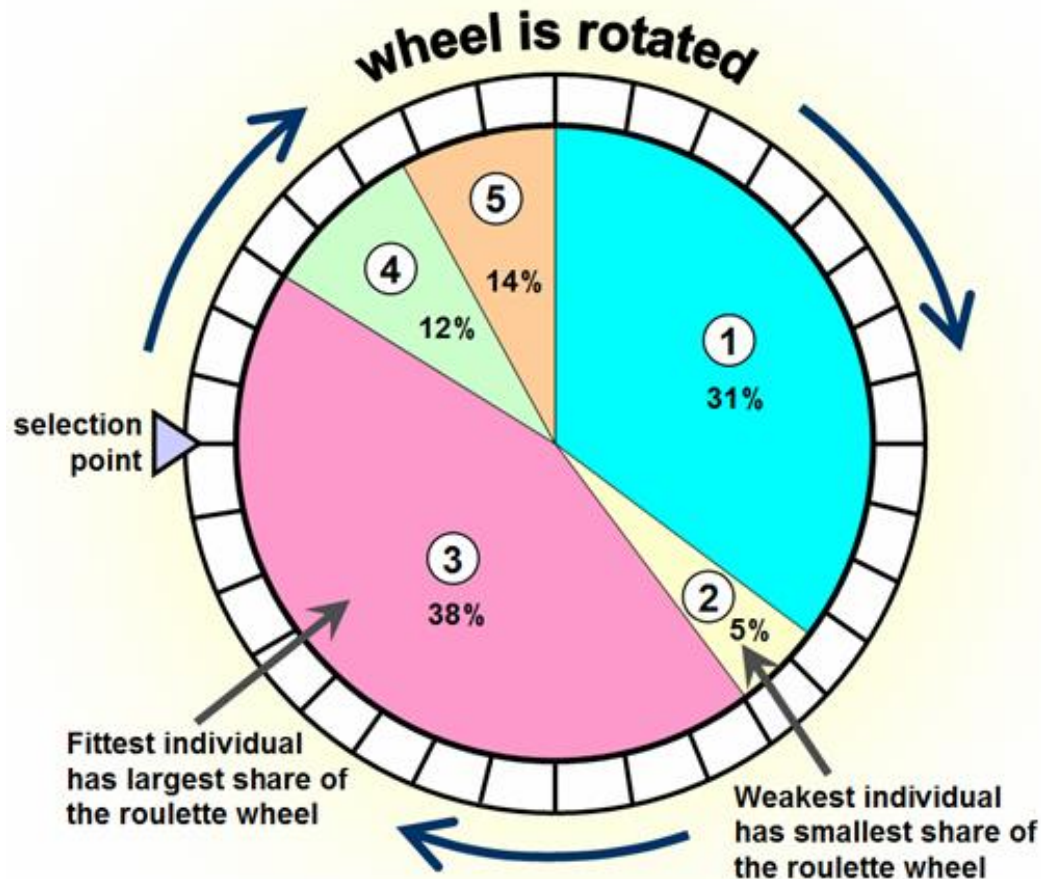
Rank Based Selection

- Attempts to remove problems of FPS by basing selection probabilities on *relative* rather than *absolute* fitness

- Rank population according to fitness and then base selection probabilities on rank where the fittest has rank μ and worst rank 1

- This imposes a sorting overhead on the algorithm, but this is usually negligible compared to the fitness evaluation time.

Rank Based Selection (Cont'd)

- Selection is independent of actual fitness values, with the advantage that the best individual will not dominate in the selection process.

- It preserves a constant selection pressure by sorting the population on the basis of fitness, and then allocating selection probabilities to individuals according to their rank, rather than according to their actual fitness values.

Tournament Selection

- FP and R selection methods and the algorithms used to sample from their probability distribution relied on a knowledge of the entire population.

- In certain situations, if the population is very large, or if the population is distributed in some way (perhaps on a parallel system), obtaining this knowledge is either highly time consuming or at worst impossible.

- In other cases, there might not be a universal fitness definition at all. For instance, think of an application evolving game playing strategies. In this case we might not be able to quantify the strength of a given strategy but we can compare any two of them by simulating a game played by these strategies as opponents.

Tournament Selection (Cont'd)

- Informal Procedure:
  - Pick n_{ts} members at random then select the best of these
  - Repeat to select more individuals

- Inherits the advantage of rank selection

- Does not require global reordering

- For crossover with two parents, tournament selection is done twice, once for the selection of each parent.

Tournament Selection (Cont'd)

- Provided that the tournament size, n_{ts}, is not too large, tournament selection prevents the best individual from dominating, thus having a lower selection pressure.

- On the other hand, if n_{ts} is too small, the chances that bad individuals are selected increase. Thus, the elective pressure is directly related to n_{ts}.

- *If $n_{ts} = n_s$, the best individual will always* be selected, resulting in a very high selective pressure.

- On the other hand, if $n_{ts} = 1$, random selection is obtained.

Tournament Selection (Cont'd)

- There also exists a non-deterministic variant of this selection where this is not necessarily the case. Therefore, a probability p is defined. The best individual in the tournament is selected with probability p, the second best with probability p(1 – p), the third best with probability p(1 – p)2 and so on.

- Tournament selection is perhaps the most widely used selection operators in the modern applications of EAs, due to its extreme simplicity and the fact that the selection pressure is easy to control by varying the tournament size.

Replace Worst/Truncation

- In this scheme the worst lambda members of the population are selected for replacement.

- Although this can lead to very rapid improvements in the mean population fitness, it can also lead to premature convergence as the population tends to rapidly focus on the fittest member currently present.

Random Selection

- Random selection is the simplest selection operator, where each individual has the same probability *to be selected.*

- *No fitness* information is used, which means that the best and the worst individuals have exactly the same probability of surviving to the next generation.

- Random selection has the lowest selective pressure.

Random Selection

- Random selection returns elements by chance. A possible preceding fitness assignment process as well as the objective values of the individuals play no role at all.

- This hinders the optimization algorithm to follow any gradient in the fitness landscape – it is effectively turned into a random walk.

- Random selection is thus not applied exclusively, but can serve as mating selection scheme in conjunction with a separate environmental selection.

- It maximally preserves the diversity and can be a good choice if used to pick elements from an optimal set.

Elitism

- Elitism refers to the process of ensuring that the best individuals of the current population survive to the next generation.
- The best individuals are copied to the new population without being mutated.
- The more individuals that survive to the next generation, the less the diversity of the new population.
- This scheme is commonly used in conjunction with age-based and stochastic fitness-based replacement schemes, in an attempt to prevent the loss of the current fittest member of the population.

Elitist Selection

- In elitism at least one copy of the best individual in the population is always passed onto the next generation

  – The main advantage is that convergence is guaranteed (i.e., if the global maximum is discovered, the EA converges to the maximum).

  – By the same token, however, there is a risk of being trapped in a local maximum.

Hall of Fame

- The hall of fame is a selection scheme similar to the list of best players of an arcade game.

- For each generation, the best individual is selected to be inserted into the hall of fame.

- The hall of fame will therefore contain an archive of the best individuals found from the first generation.

- The hall of fame can be used as a parent pool for the crossover operator, or, at the last generation, the best individual is selected as the best one in the hall of fame.

Selective Pressure

- Selection operators are characterized by their selective pressure, also referred to as the takeover time, which relates to the time it requires to produce a uniform population.

- It is defined as the speed at which the best solution will occupy the entire population by repeated application of the selection operator alone.

- An operator with a high selective pressure decreases diversity in the population more rapidly than operators with a low selective pressure, which may lead to premature convergence to suboptimal solutions.

- A high selective pressure limits the exploration abilities of the population.

Selective Pressure(Cont'd)

- Schemes based on exploration are said to have a low selection pressure, while schemes based on exploitation are said to have greater selection pressure.

- In other words, selection pressure is a vague measure of how often more fit individuals are selected to reproduce and/or live to the next generation.

Selection Procedure (Cont'd)

- Selection schemes can be further categorized into generational or steady-state schemes.

- A selection scheme is ***generational*** when the entire current population is replaced by its offspring to create the next generation

- A scheme is referred to as ***steady-state*** when a selected few offspring replace a few members of the current generation to form the next generation.

Survivor Selection

- a.k.a. ***replacement***
- Most EAs use fixed population size so need a way of going from (parents + offspring) to next generation
- Often deterministic
  - Fitness based : e.g., rank parents+offspring and take best
  - Age based: make as many offspring as parents and delete all parents
- Sometimes do combination (elitism)

# Age-Based and Fitness-based Replacement

- The basis of age-based schemes is that the fitness of individuals is not taken into account during the selection of which individuals to replace in the population, rather they are designed so that each individual exists in the population for the same number of EA iterations.

- For fitness-based replacement, a wide number of strategies have been proposed for choosing which μ of the $\mu + \lambda$ parents and offspring should go forward to the next EA iteration.

Performance Indicator

- **Best-so-far (BSF) -** We record the best solution found by the algorithm thus far for each generation in every run.

- **Average-of-current-population (ACP) -** We record the average solution in each generation in every run.

- **Worst-of-current-population (WCP) -** We record the worst solution in each generation in every run.
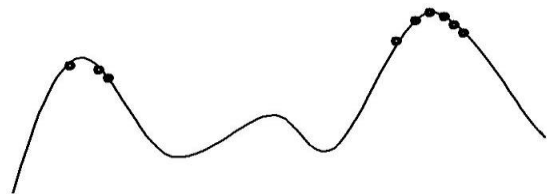
Typical behaviour of an EA

Phases in optimising on a 1-dimensional fitness landscape
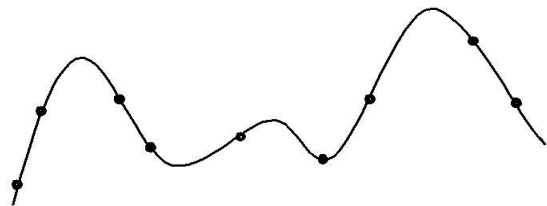
Early phase:

quasi-random population distribution
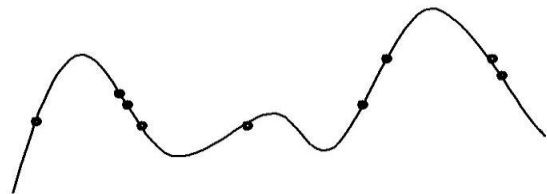
Mid-phase:

population arranged around/on hills

Late phase:

population concentrated on high hills

Thanks