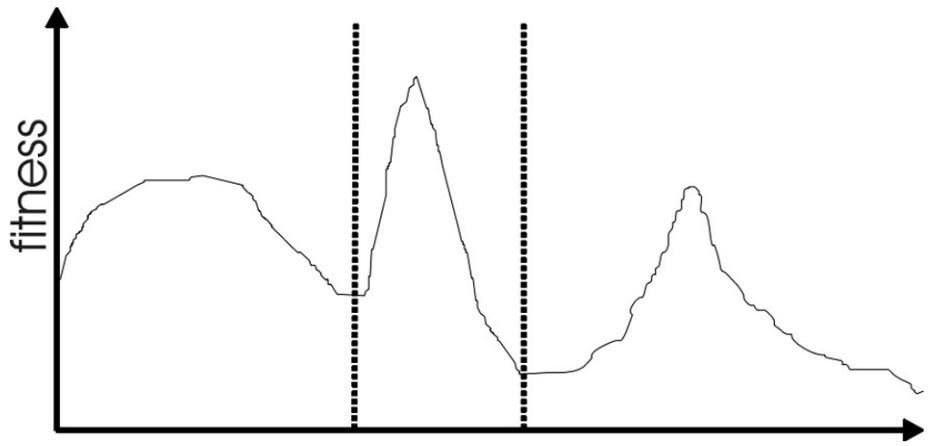
9: Selection and the Need for Diversity

- Selection pressure and diversity
- Multi-modal optimization
- Techniques to preserve diversity
  - Crowding
  - Automatic speciation
  - Multi-population island model
  - Cellular EAs
  - Quality-diversity optimization
- Textbook Chapter 5.5

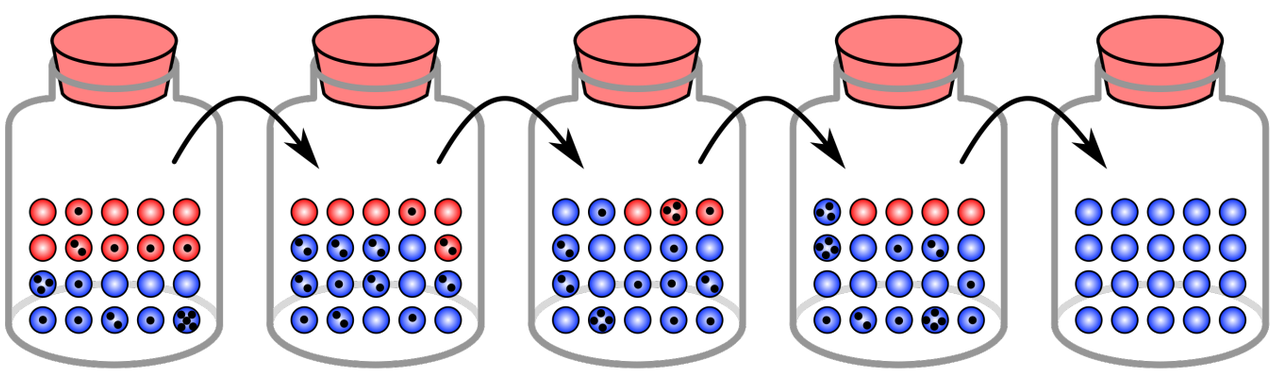
Selection pressure and diversity

- Selection pressure
  - the pressure of pushing quality improvement
  - higher selection pressure -> fitter solutions are more likely to be chosen (parent or survivor)
  - population size, tournament size, fitness or ranking scaling
- Diversity
  - higher selection pressure leads to lower diversity
  - diversity of genotypes, phenotypes, and fitness
  - maintain diversity is important for multimodal problems

Multimodality



- Most interesting and practical problems have more than one locally optimal solutions
- Often useful to identify several possible peaks (options)
- Sub-optima can be useful sometimes
- EAs are good at identifying multiple optima
- Genetic drift - one optimum eventually takes over the entire population



Preserving diversity

- Explicit approaches
  - make similar individuals compete for resources (fitness)
  - make similar individuals compete with each other for survival
- Implicit approaches
  - impose an equivalent of geographical separation
  - impose an equivalent of speciation
- Different spaces
  - genotype space
  - phenotype space

Fitness sharing

- Idea
  - control the number of individuals within a niche by sharing their fitness
  - the shared fitness is computed as the original one adjusted according to the number of individuals in the niche
  - need to compute the pairwise distance (phenotypically or genotypically) and set a distance threshold to define niches
- Implementation
  - set the size of the niche using a distance threshold σ in phenotype/genotype space
  - adjust the fitness of individual i

$$f'(i) = \frac{f(i)}{\sum_{j=1}^{\mu} sh(d(i,j))} \qquad sh(d) = \begin{cases} 1 - d/\sigma & d \le \sigma \\ 0 & otherwise \end{cases}$$

Crowding

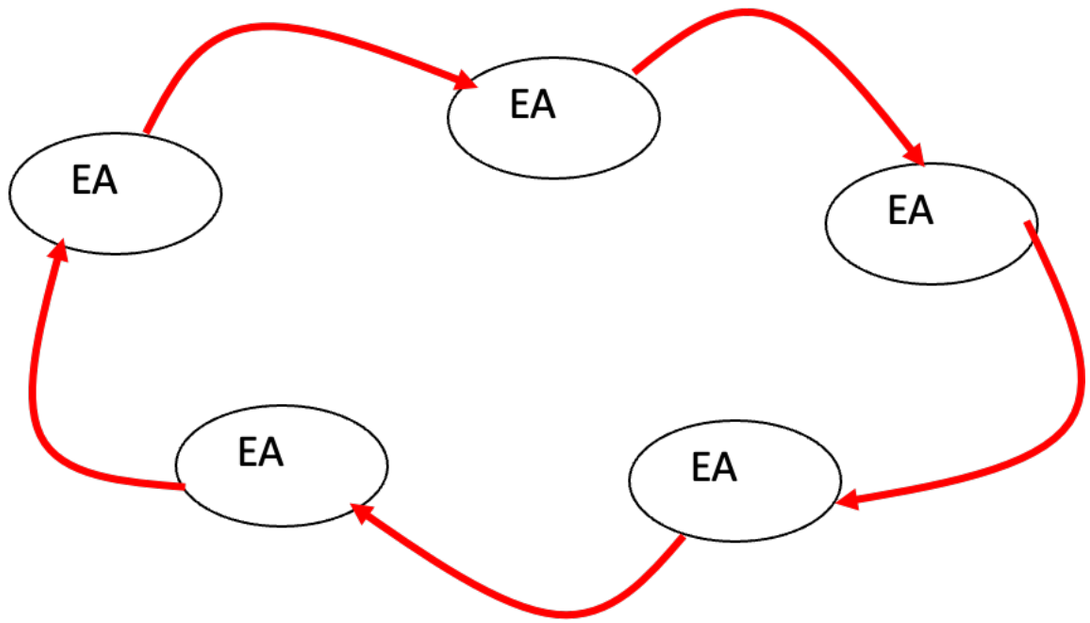
- Idea
  - ensure that new individuals replace similar members of the population
  - the offspring replaces the most similar of the parents
  - attempt to distribute individuals evenly among niches
- Implementation
  - use a distance metric in the phenotype/genotype space
  - randomly shuffle and pair parents, produce offspring (e.g., 20% of the population size)
  - a set of members of the parent population are chosen at random
  - each offspring replaces the most similar of those older generation
  - or let offspring compete directly with parents (assume that parents and offspring are similar)

Automatic speciation

- Idea
  - impose mating restrictions based on some aspect of the candidate solutions defining them as belonging to different species
  - only mate with members from the same species
- Implementation
  - use a pairwise distance metric in phenotype/genotype space to define species
  - only allow recombination and selection within species

Multi-population island model

- Idea
  - evolve multiple populations in parallel
  - allow migration from time to time
- Implementation
  - run multiple EAs in parallel
  - after a fixed number of generations (an epoch), exchange some individuals with neighboring population

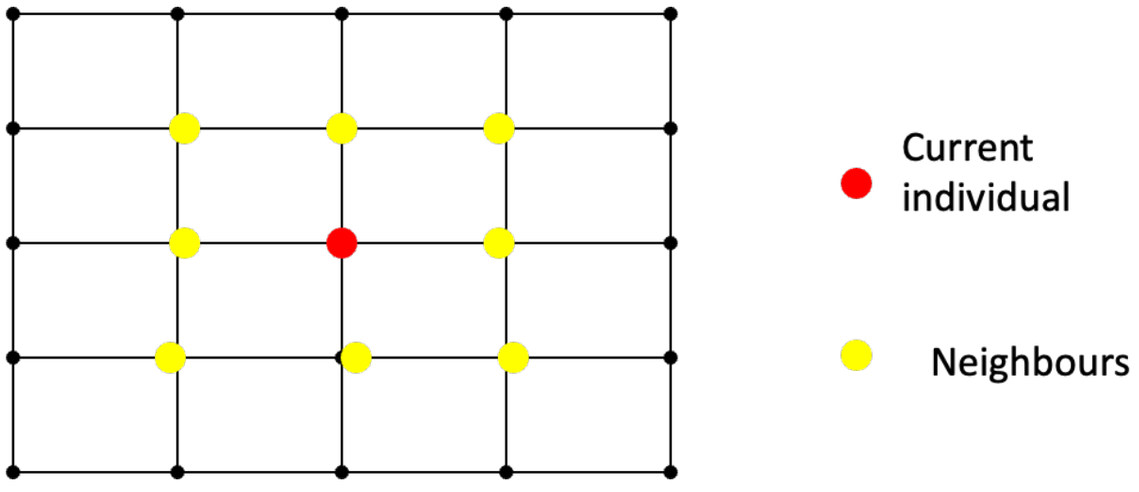


Island model parameters

- How often to exchange individuals?
  - too frequently, essentially running in one population
  - too infrequently, some populations may have converged already
- How many individuals to exchange?
  - to prevent rapid spread of super individuals
  - exchange a small number of solutions at a time, e.g., 2 to 5
- Which individuals to exchange?
  - fitness based or random
- How many populations?
  - usually more is beneficial if computational resources allow

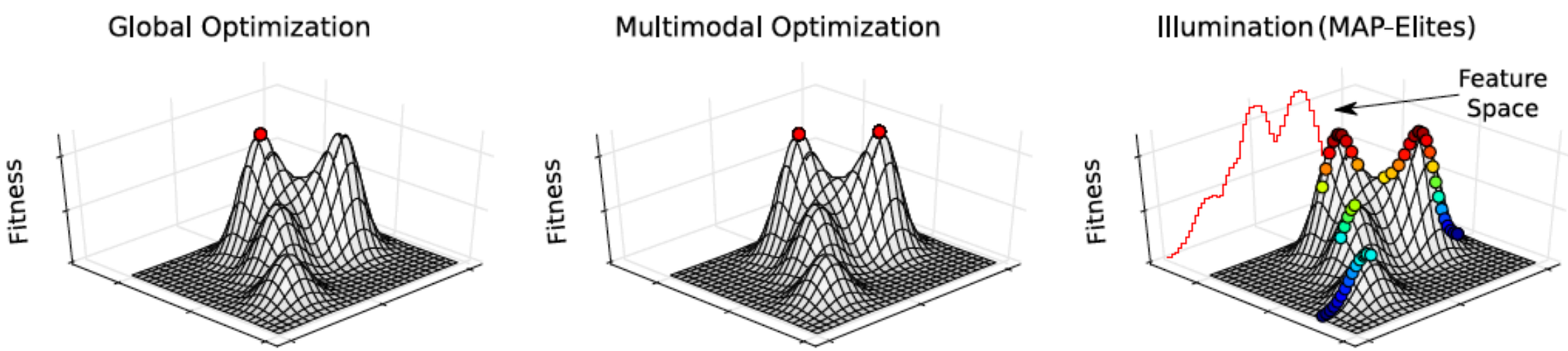
Spatial distribution - cellular EAs

- Idea
  - a single population split into a large number of smaller overlapping subgroups
  - only allow recombination within neighborhood
  - similar to biological systems
- Implementation
  - each member of the population exists at a different point on a grid
  - each node has a defined deme (neighborhood), usually same size for all nodes
  - in each generation, consider each deme for selection, recombination, mutation, and replacement



Quality-diversity optimization

- Idea
  - explicitly designed to search for a large collection of both diverse and high-performing solutions
- Implementation
  - define descriptors for solutions
  - map individuals onto the descriptor space (could be high dimensional)
  - divide the space into cells with fixed number of individuals
  - only allow individuals in the same cell to compete and recombine



Vassiliades et al., IEEE Transactions on Evolutionary Computing, 22:4, 623-630, 2018