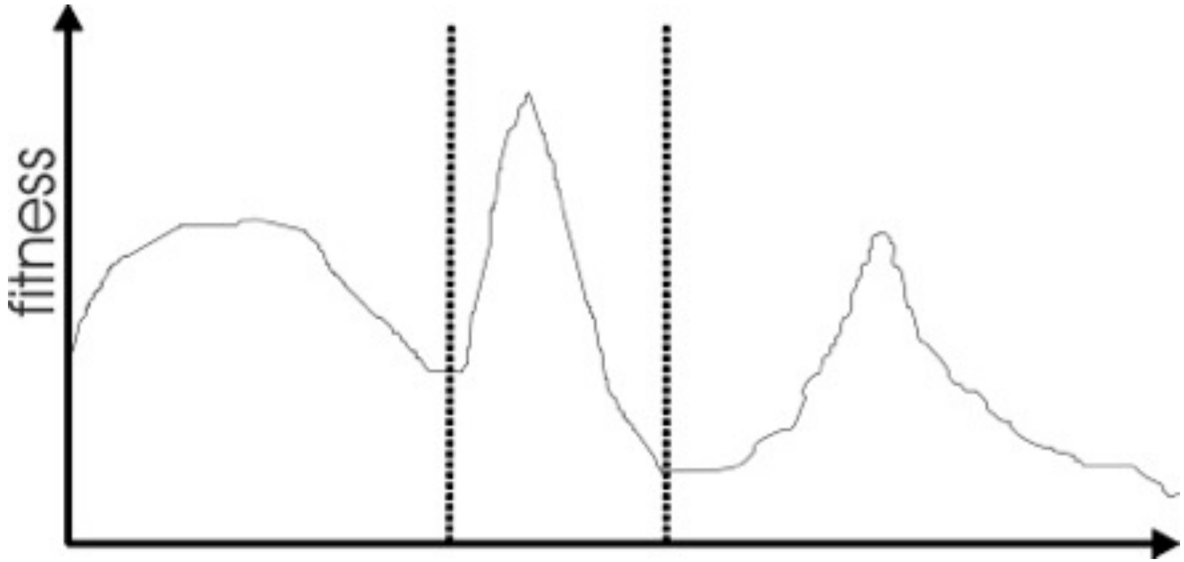
Multimodal Problems and Spatial Distribution

Lecture 10

MONASH UNIVERSITY CLAYTON'S SCHOOL OF INFORMATION TECHNOLOGY

Motivation 1: Multimodality

Reading: Eiben& Smith Chapter 9



Most interesting problems have more than one locally optimal solution.

Motivation 2: Genetic Drift

- Finite population with global (panmictic) mixing and selection eventually convergence around one optimum
- Often might want to identify several possible peaks
- This can aid global optimisation when sub-optima has the largest basin of attraction

Biological Motivation 1: Speciation

- In nature different species adapt to occupy different environmental niches, which contain finite resources, so the individuals are in competition with each other
- Species only reproduce with other members of the same species (Mating Restriction)
- These forces tend to lead to phenotypic homogeneity within species, but differences between species

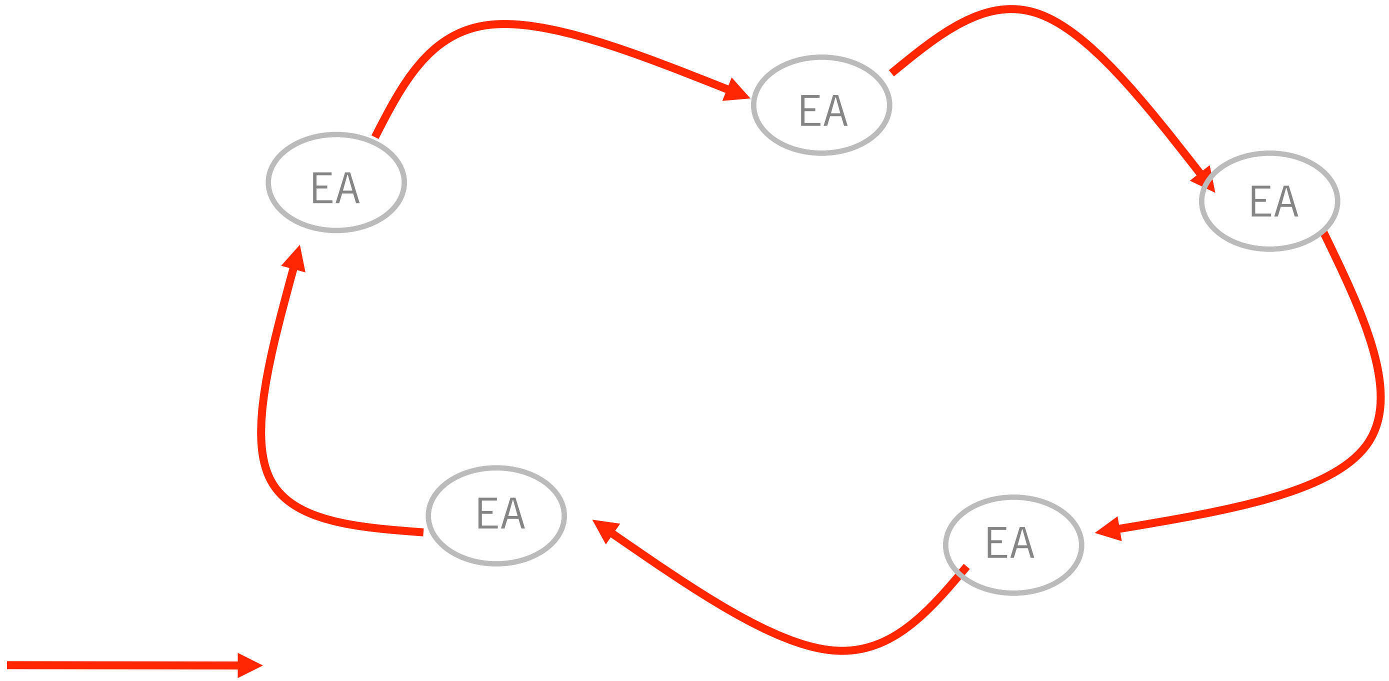
Biological Motivation 2: Punctuated Equilbria

- Theory that periods of stasis are interrupted by rapid growth when main population is "invaded" by individuals from previously spatially isolated group of individuals from the same species
- The separated sub-populations (demes) often show local adaptations in response to slight changes in their local environments

Implications for Evolutionary Optimisation

- Two main approaches to diversity maintenance:
- **Implicit approaches**
  - Impose an equivalent of geographical separation
  - Impose an equivalent of speciation
- **Explicit approaches**
  - Make similar individuals compete for resources (fitness)
  - Make similar individuals compete with each other for survival

Implicit 1: "Island" Model Parallel EAs



Periodic migration of individual solutions between populations

Island Model EAs contd:

- Run multiple populations in parallel, in some kind of communication structure (usually a ring or a torus, possibly hypercube).
- After a (usually fixed) number of generations (an **Epoch**), exchange individuals with neighbours
- Repeat until ending criteria met
- Partially inspired by parallel/clustered systems

Island Model Parameters 1

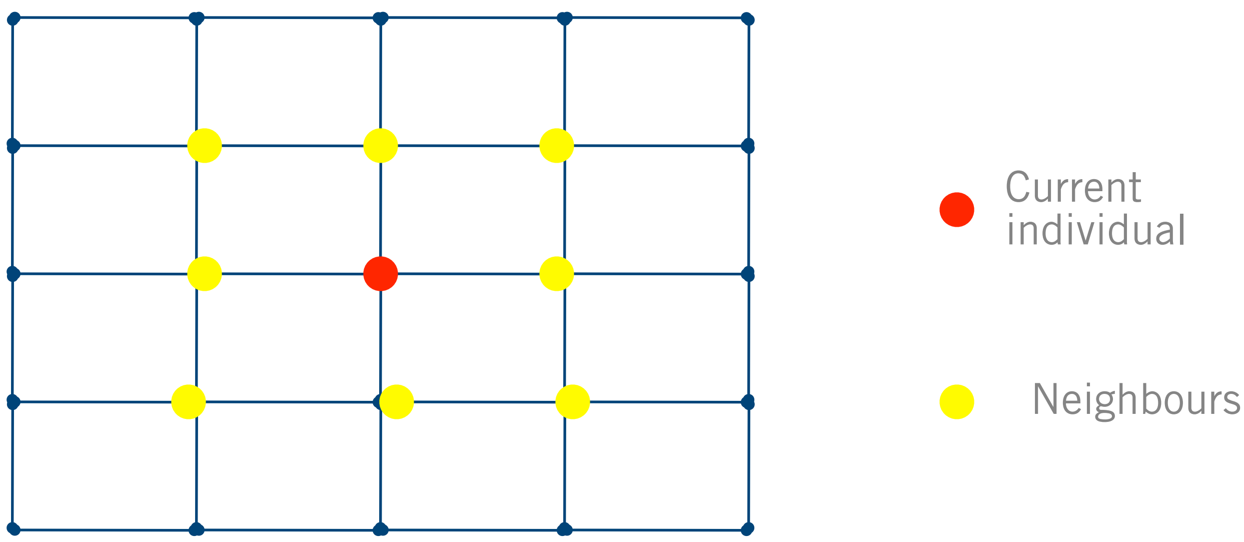
- Could use different operators in each island
- How often to exchange individuals ?
  - too quick and all pops converge to same solution
  - too slow and waste time
  - most authors use range~ 25-150 gens
  - can do it adaptively (stop each pop when no improvement for (say) 25 generations)

Island Model Parameters 2

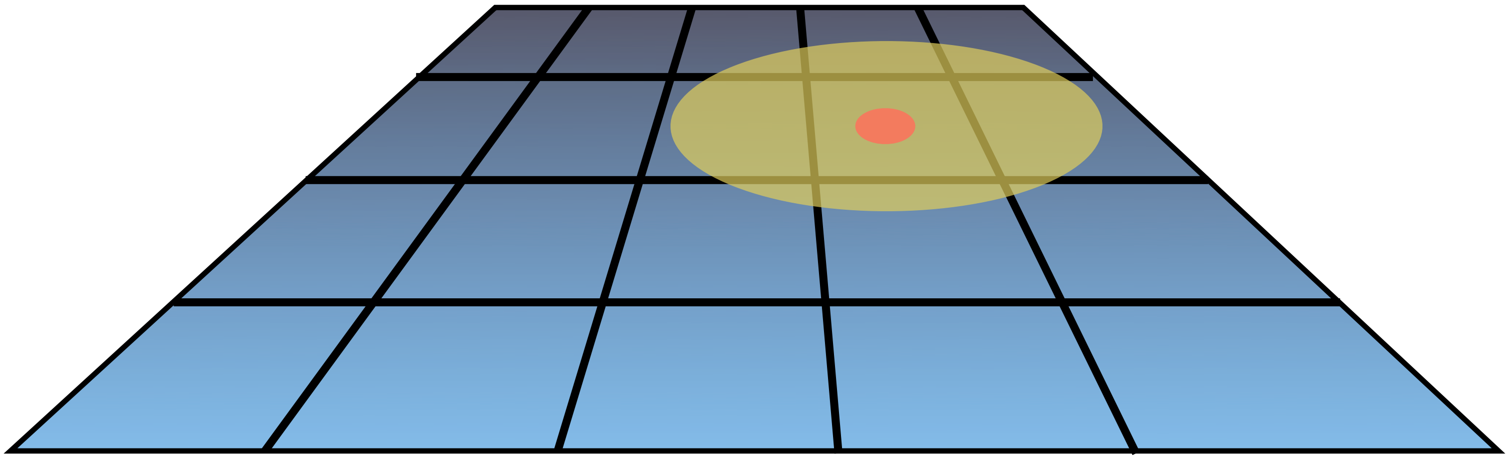
- How many, which individuals to exchange ?
  - usually ~2-5, but depends on population size.
  - more sub populations usually gives better results but there can be a "critical mass" i.e. minimum size of each sub population needed
  - Martin et al found that better to exchange randomly selected individuals than best
  - can select random/worst individuals to replace

Implicit 2: Diffusion Model Parallel EAs

- Impose spatial structure (usually grid) in 1 pop



Geographically Distributed EAs



- Distribute population over a 2D grid.
- Local Selection
- Asynchronous
- Parallelisation possible

Geographically Distributed EAs (cont.)

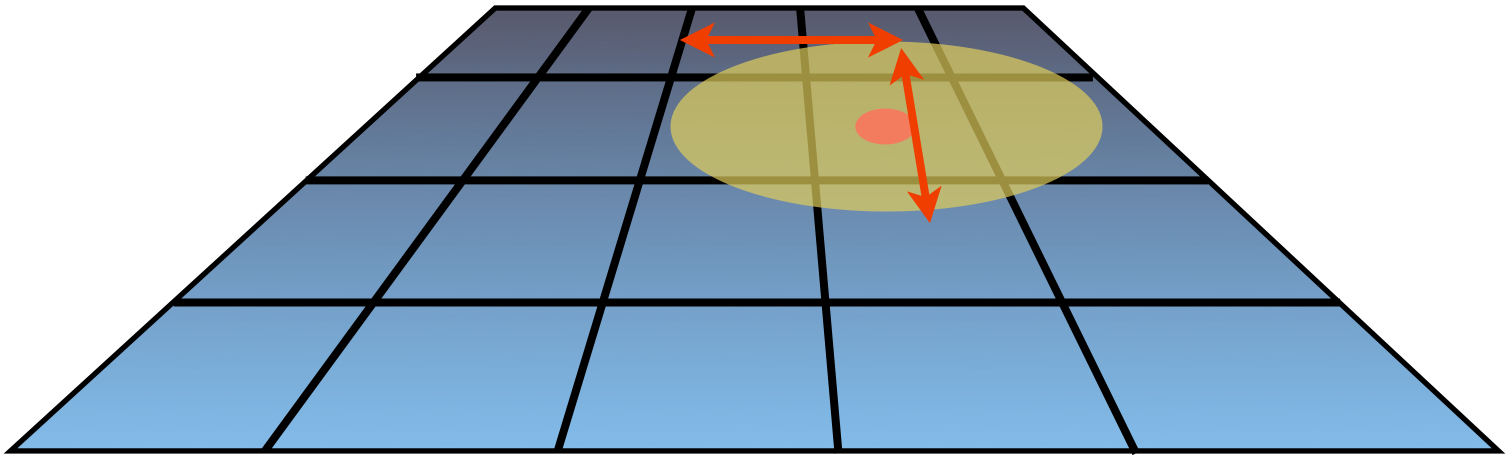
1. Create random genotypes at each cell on a 2D toroidal grid;
2. Randomly pick cell on grid, C, this holds genotype Cg
3. Create a set of cells, S, in the neighbourhood of C
4. Select (proportional to fitness) a genotype, m, from one of the cells in S
5. Create offspring O, from m and Cg
6. Select (inversely proportional to fitness) a genotype d at one of the cells in S
7. Replace d with O
8. Goto step 2

Diffusion Model EAs

- Consider each individual to exist on a point on a (usually rectangular toroid) grid
- Selection (hence recombination) and replacement happen using concept of a neighbourhood a.k.a. **deme**
- Leads to different parts of grid searching different parts of space, good solutions diffuse across grid over a number of generations

Creating a Neighbourhood

1. Choose Δx, Δy from Gaussian probability distribution, flip whether +/– direction



2. define sets of cells at distance 1,2,3... from current cell. Pick distance from Gaussian distribution, pick cell at this distance randomly
3. *N* random walks
4. Deterministic (e.g. 8 nearest neighbours)

Diffusion Model Example

- Assume rectangular grid so each individual has 8 immediate neighbours
- equivalent of 1 generation is:
  - pick point in pop at random
  - pick one of its neighbours using roulette wheel
  - crossover to produce 1 child, mutate
  - replace individual if fitter
  - circle through population until done

Implicit 3: Automatic Speciation

- Either only mate with genotypically / phenotypically similar members or
- Add bits to problem representation
  - that are initially randomly set
  - subject to recombination and mutation
  - when selecting partner for recombination, only pick members with a good match
  - can also use tags to perform fitness sharing (see later) to try and distribute members amongst niches

Explicit 1: Fitness Sharing

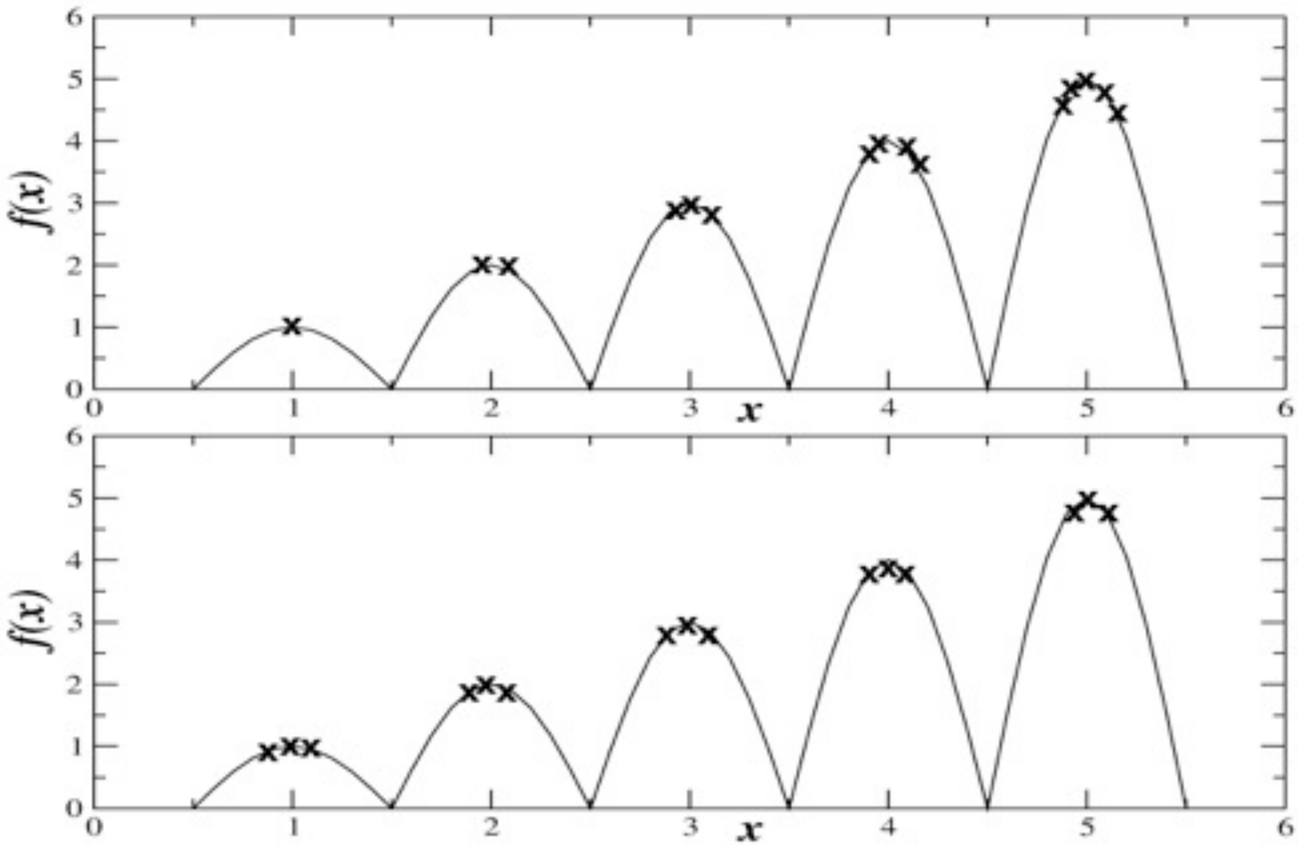
- Restricts the number of individuals within a given niche by "sharing" their fitness, so as to allocate individuals to niches in proportion to the niche fitness
- need to set the size of the niche σ_{share} in either genotype or phenotype space
- run EA as normal but after each gen set

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

Explicit 2: Crowding

- Attempts to distribute individuals evenly amongst niches
- relies on the assumption that offspring will tend to be close to parents
- uses a distance metric in ph/g enotype space
- randomly shuffle and pair parents, produce 2 offspring
- 2 parent/offspring tournaments - pair so that d(p1,o1)+d(p2,o2) < d(p1,02) + d(p2,o1)

Fitness Sharing vs. Crowding



Multi-Objective Problems (MOPs)

- Wide range of problems can be categorised by the presence of a number of n possibly conflicting objectives:
  - buying a car: speed vs. price vs. reliability
  - engineering design: lightness vs strength
- Two part problem:
  - finding set of good solutions
  - choice of best for particular application

MOPs 1: Conventional approaches

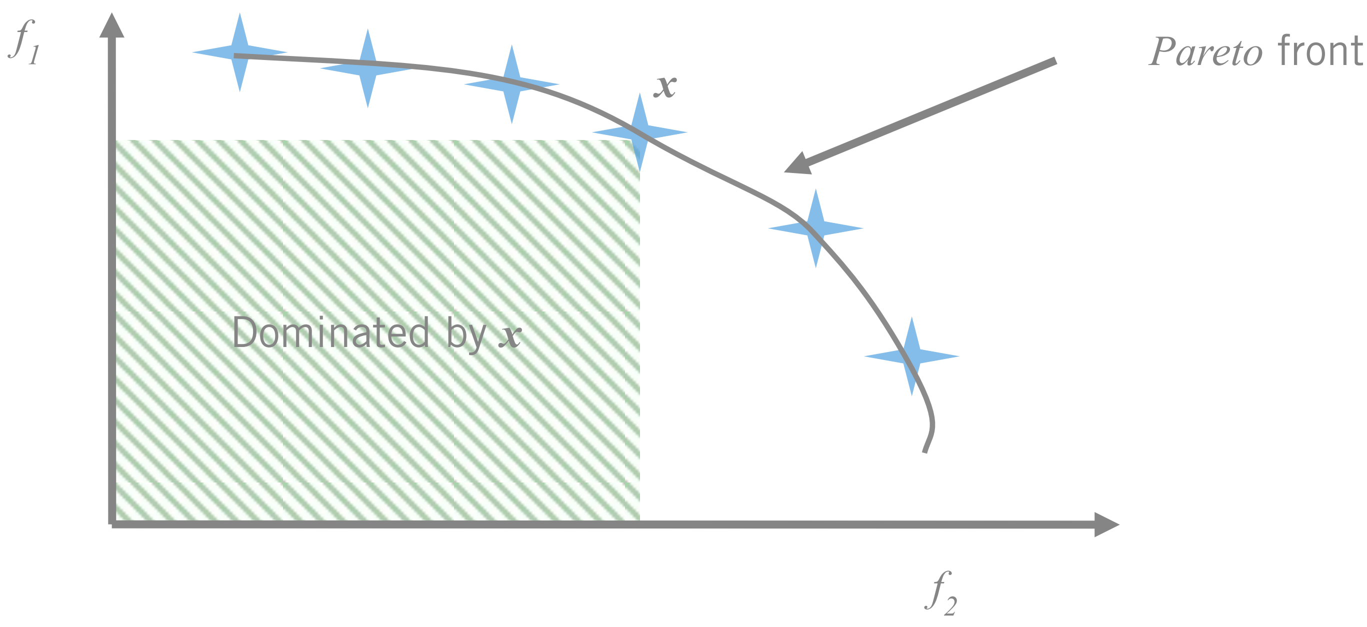
- rely on using a weighting of objective function values to give a single scalar objective function which can then be optimised:

$$f'(x) = \sum_{i=1}^{n} w_i f_i(x)$$

- to find other solutions have to re-optimise with different w_i.

MOPs 2: Dominance

- we say x dominates y if it is at least as good on all criteria and **better** on at least one



MOPs 3: Advantages of EC approach

- Population-based nature of search means you can *simultaneously* search for set of points approximating Pareto front
- Don't have to make guesses about which combinations of weights might be useful
- Makes no assumptions about shape of Pareto front - can be convex / discontinuous etc

MOPs 4: Requirements of EC approach

- Way of assigning fitness,
  - usually based on dominance
- Preservation of diverse set of points
  - similarities to multi-modal problems
- Remembering all the non-dominated points you've seen
  - usually using elitism or an archive

MOPs 5: Fitness Assignment

- Could use aggregating approach and change weights during evolution
  - no guarantees
- Different parts of pop use different criteria
  - e.g. VEGA, but no guarantee of diversity
- Dominance
  - ranking or depth based
  - fitness related to whole population

MOPs 6: Diversity Maintenance

- Usually done by niching techniques such as:
  - fitness sharing
  - adding amount to fitness based on inverse distance to nearest neighbour (minimisation)
  - (adaptively) dividing search space into boxes and counting occupancy
- All rely on some distance metric in genotype / phenotype space

MOPs 7: Remembering Good Points

- Could just use elitist algorithm
  - e.g. ($\mu + \lambda$) replacement
- Common to maintain an archive of non-dominated points
  - some algorithms use this as second population that can be in recombination etc.
  - others divide archive into regions too e.g. PAES