

Fitness Landscapes

Seth Bullock

bristol.ac.uk

This time:

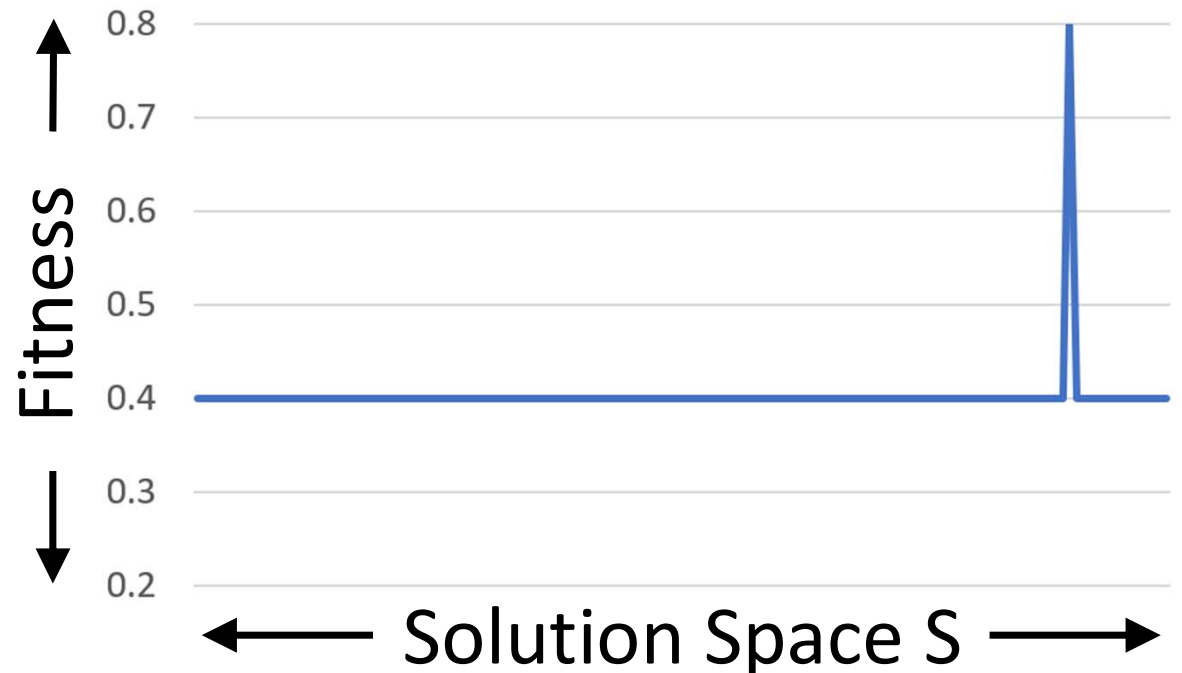
- Visualizing search problems as fitness landscapes
- Kinds of landscape (unimodal, multimodal, deceptive...)
- How search operators interact with fitness functions
- How search moves across landscapes
- The need to both explore new areas and exploit existing knowledge while searching

- The set of possible genotypes is the GA's search space, S
 - Given genotypes of length L from an alphabet of size A :
 - The space has L dimensions which each have A possible values
 - There are therefore A^L possible genotypes in S
 - Each genotype is a *point* in the search space
 - Genotypes are 'neighbours' if they differ by one mutation
 - The *structure* of a search space influences how challenging it is for an algorithm to search it
-

- A search space can be visualized as a *fitness landscape*
 - A plot of the fitness of all possible genotypes: fitness => height
 - Called a landscape because it can have peaks, valleys, ridges...
 - Introduced by biologist Sewall Wright in 1932
- We can think of search as movement across the landscape
 - A population of points move in discrete steps and jumps
 - The shape of the landscape influences this movement

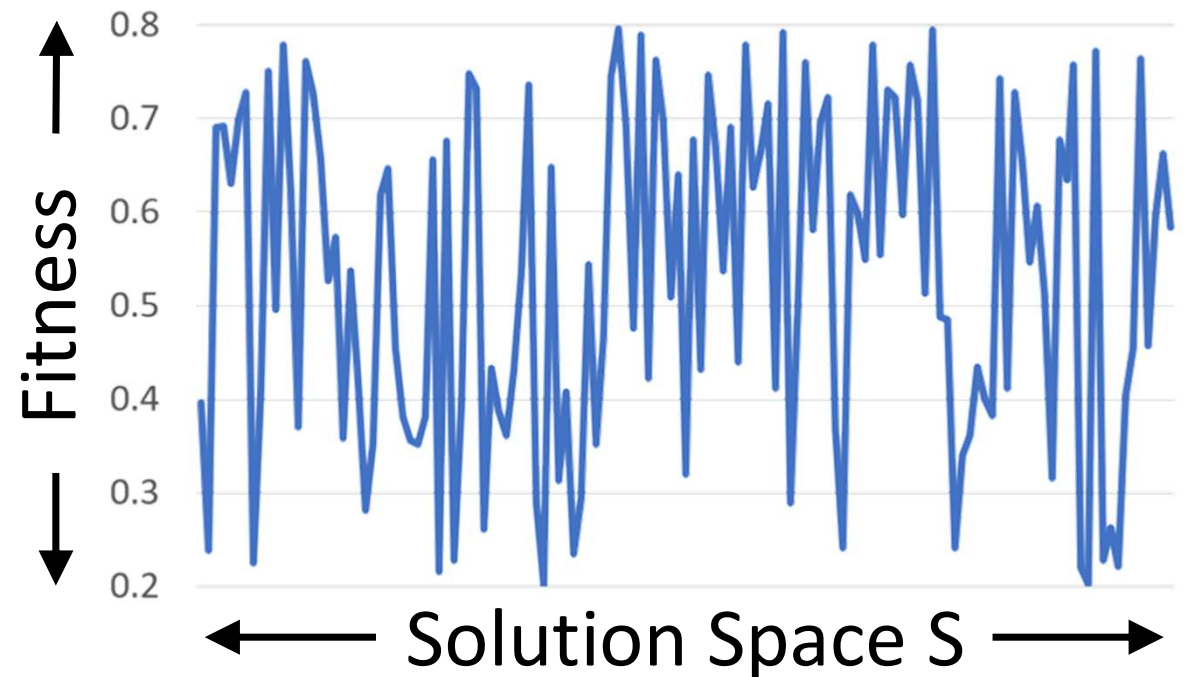
Needle-In-A-Haystack Landscape

- Suppose all genotypes except one have the *same* low fitness
- This kind of problem is *hard* for any search algorithm:
 - The fitness function is very unhelpful; it has no structure to exploit
 - The fitness of suboptimal genotypes tells the algorithm no information about where the optimal genotype is



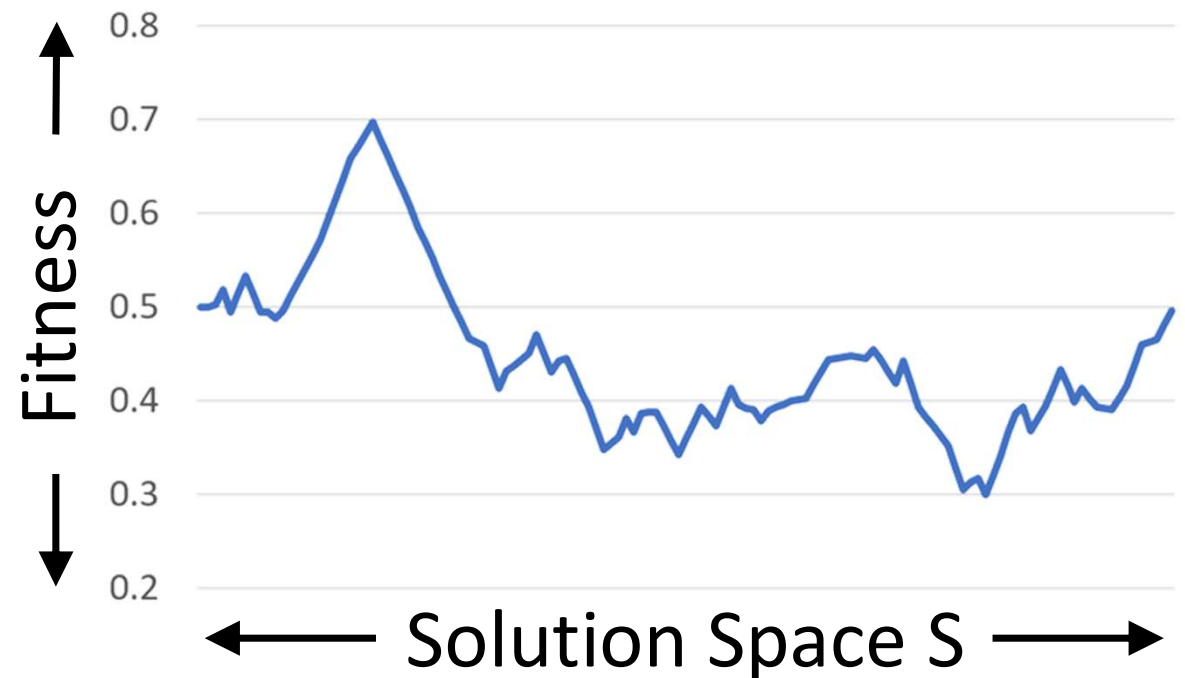
Random Landscape

- Suppose all genotypes had a *randomly* assigned fitness
- Again, this problem is *hard* for any search algorithm:
 - The fitness function is still very unhelpful; it still has no structure to exploit
 - The fitness of suboptimal genotypes still tell the algorithm no information about the optimal genotype



A Nicer Landscape

- Luckily, real-world problems do tend to have useful structure
- The algorithm's job is to be able to exploit this structure:
 - Local smoothness, i.e., nearby points have similar fitness = *correlation structure*
 - Ideally, higher peaks are also broader = *non-deceptive*
 - The problem can be divided into parts = *decomposable*



Search Neighbourhoods

- What does ‘local’ mean in the term ‘local smoothness’?
 - It depends on how we represent and search the space, S
- We could say two points, a and b , in S are ‘local’ to each other if the ‘distance’ between their genotypes is beneath a threshold
 - What distance measure to use? Hamming distance? Euclidian distance?
- Alternatively we could define the ‘distance’ between a and b in terms of how easy it is for the algorithm to move between them.
 - a and b are *neighbours* if a can reach b in one genetic operation
 - (e.g., a single mutation)

Thank you!