

Multiple Objectives – The MO in MOGA and MOGP

Greg Rohling

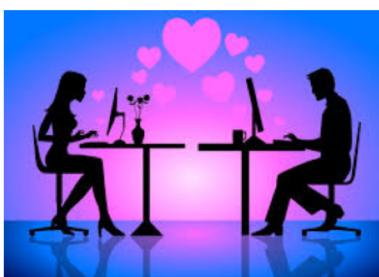
Objectives

- After this lecture students:
 - Recognize the power of multiple objective optimization in supplying a population of solutions not just a single objectives
 - Understand how Pareto dominance can be used to affect probability of mating
 - Understand classification terms true positive, false positive, true negative, false negative, sensitivity, specificity, accuracy
 - Use Multiple Objective concepts for selection of teams for project efforts

Rate Yourself

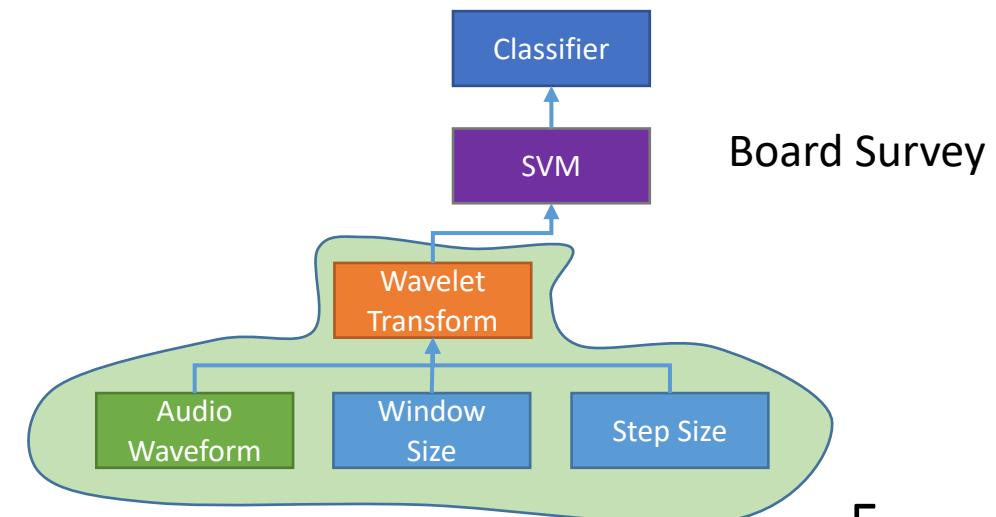
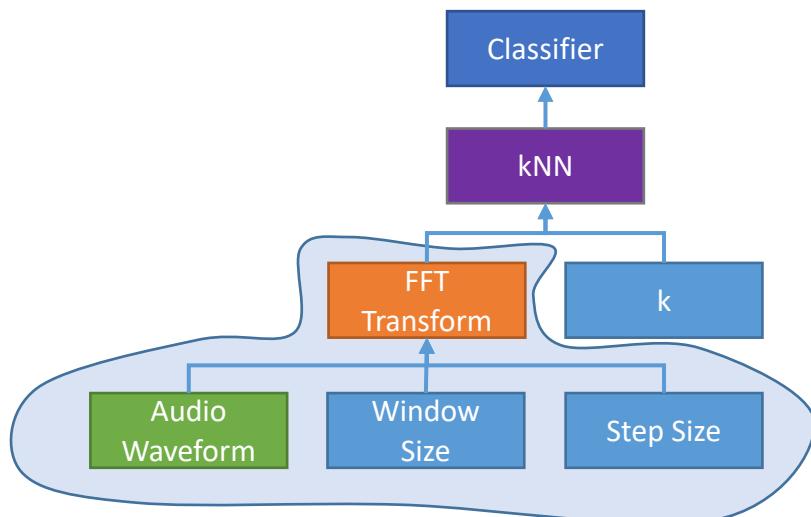
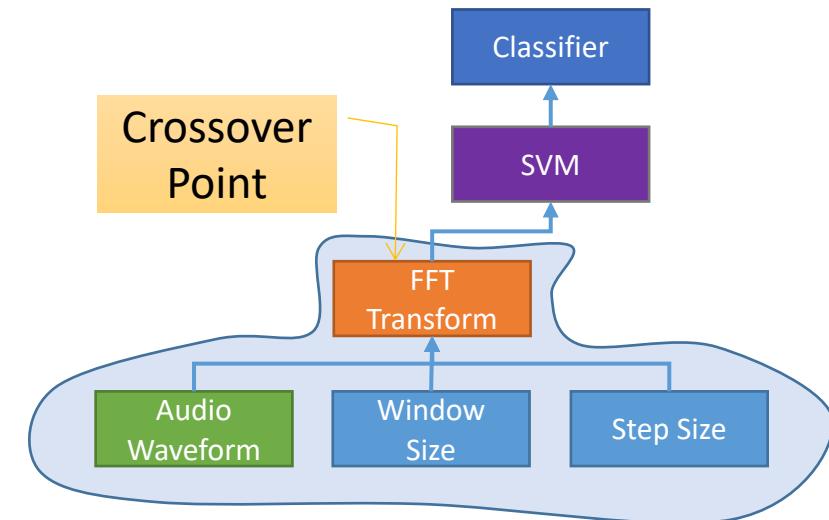
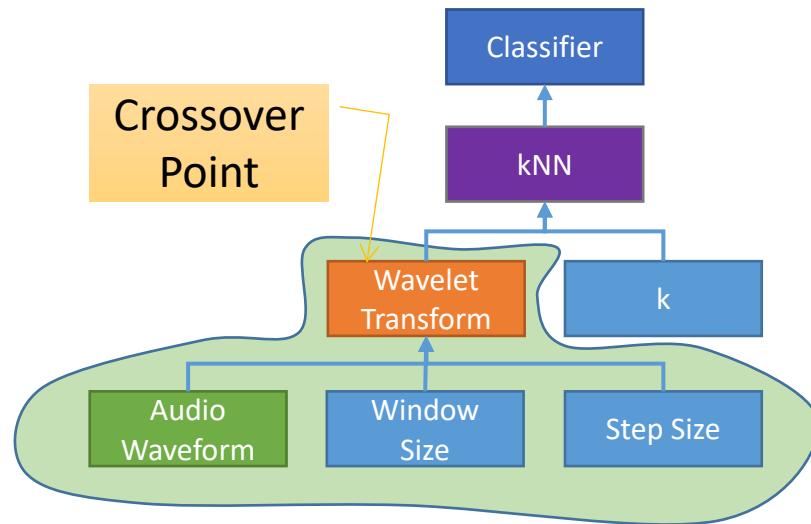
- Paper going around with 3 Columns
 - Write your name
 - Rate your python skills 0 to 9
 - 0 = none - I still think it is a snake
 - 3 = some experience
 - 7 = multiple years, and multiple applications
 - 9 = experienced and know pandas and sverkit learn
 - Rate your machine learning skills
 - 0 = none
 - 3 = read some articles, but have never used it
 - 7 = used multiple types of machine learnings on multiple applications
 - 9 = used multiple types of machine learnings on multiple applications and understand theory
- This will serve as attendance for today, and an exercise next week

What are you looking for in a date/mate?

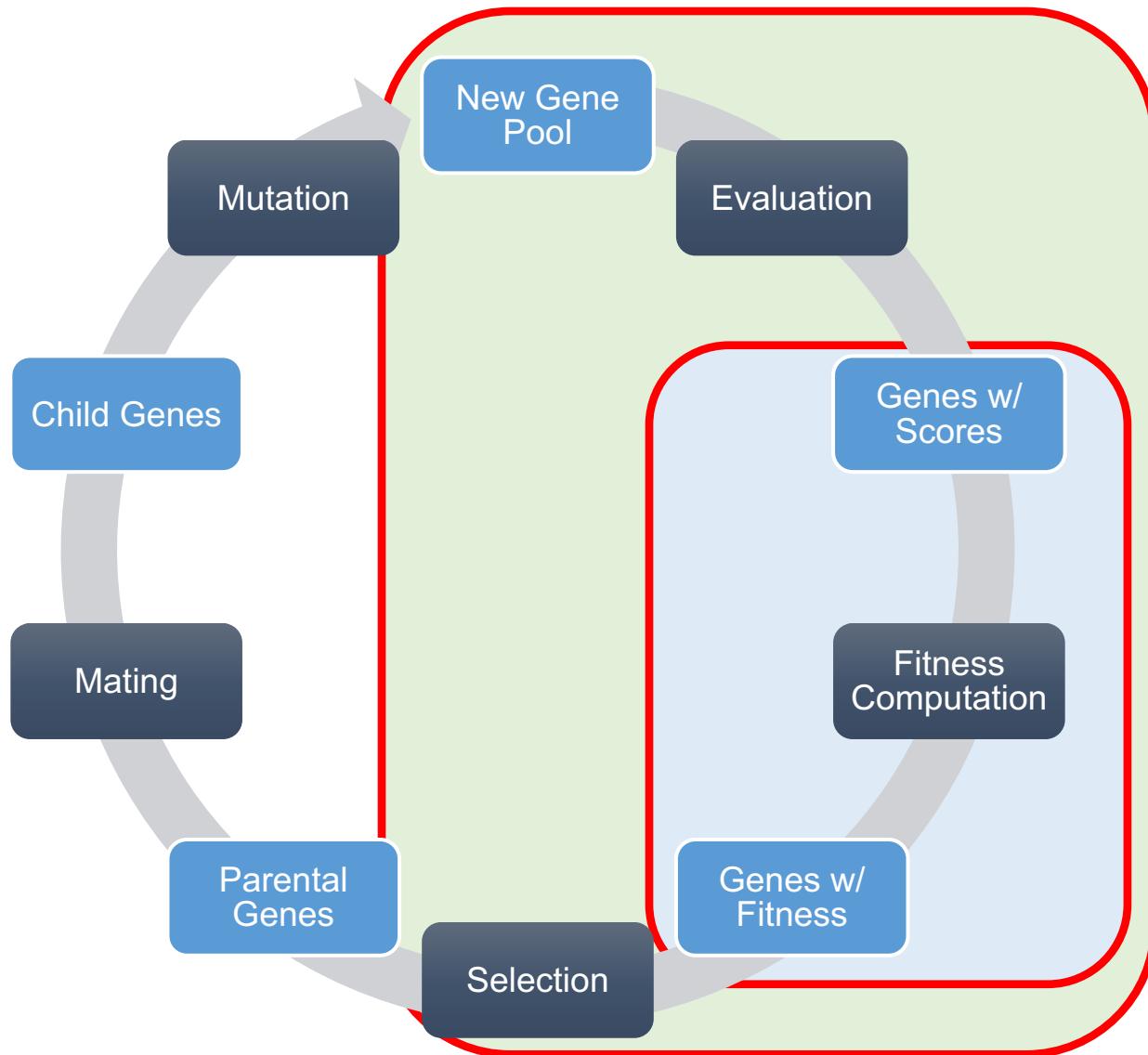


Board Survey

What is an algorithm looking for in a mate?



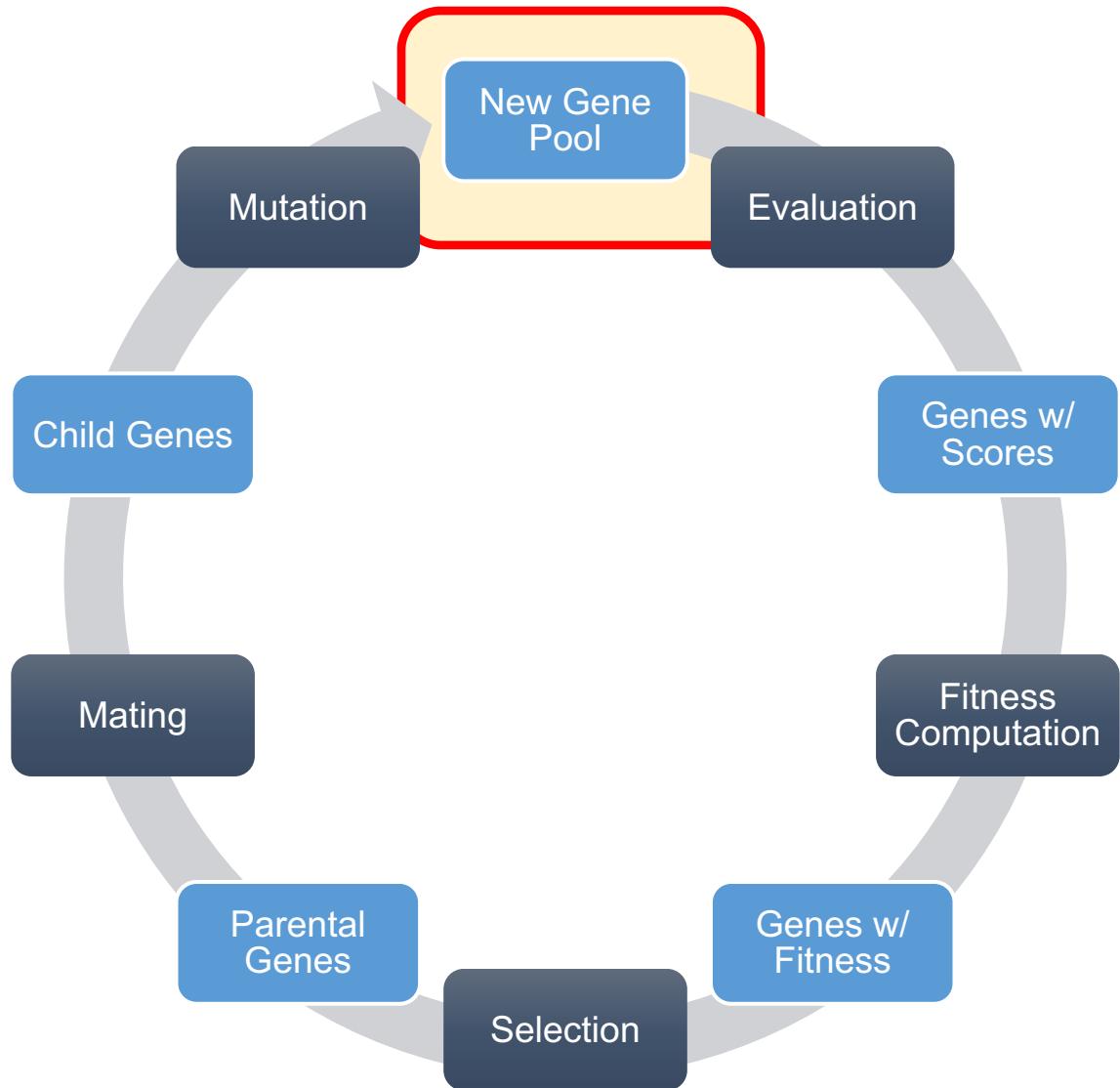
Lecture Scope



- Today's lecture focuses the translation of a vector of scores from evaluation into a fitness value

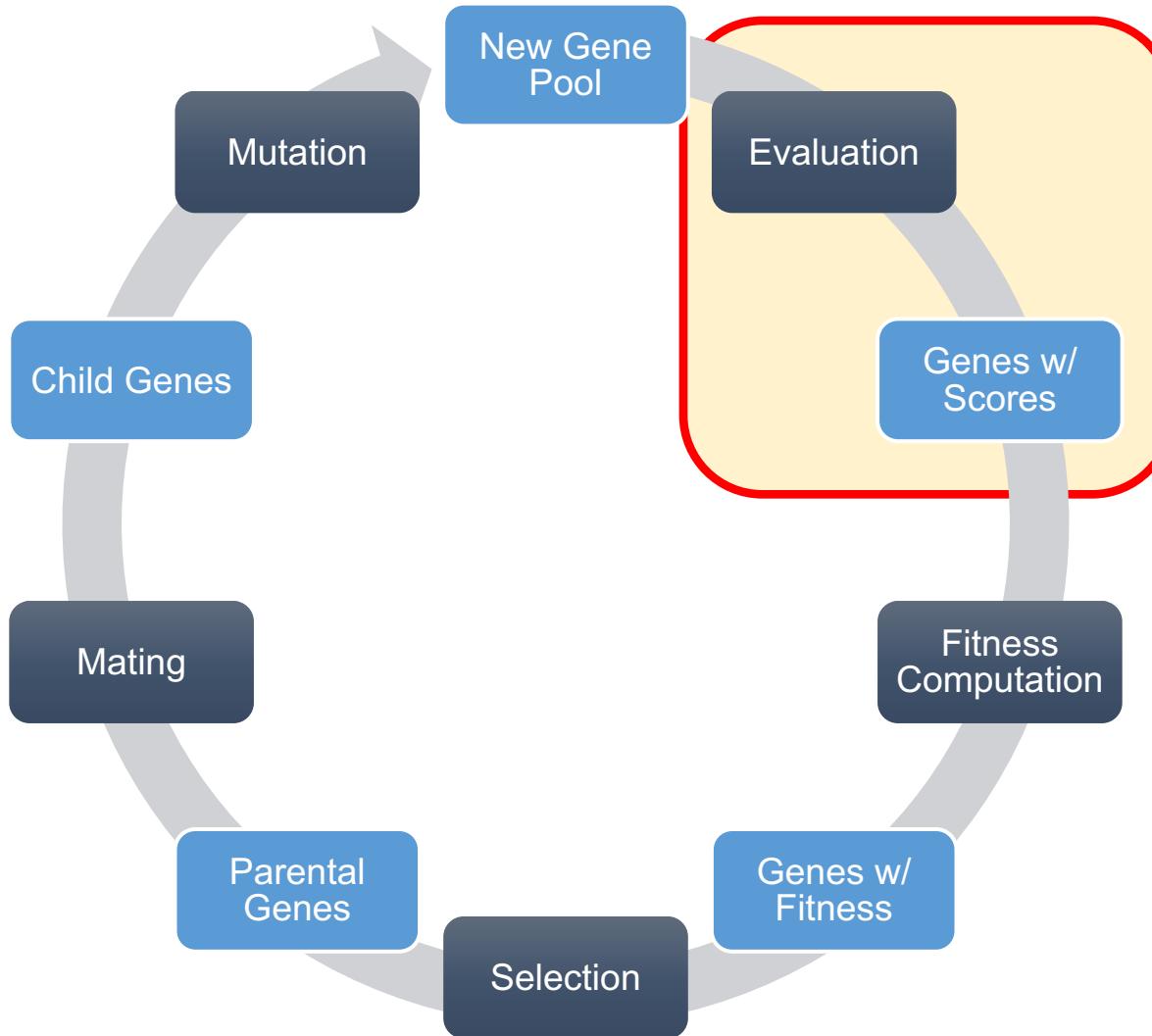
- But to understand the context we will look at the slightly larger picture

Lecture Scope



- Gene pool is the set of genome to be evaluated during the current generation
 - Genome
 - Genotypic description of an individuals
 - DNA
 - GA = set of values
 - GP = tree structure, string
 - Search Space
 - Set of all possible genome
 - For Automated Algorithm Design
 - Set of all possible algorithms
 - How big is the search space?
 - What is this important for algorithm design?

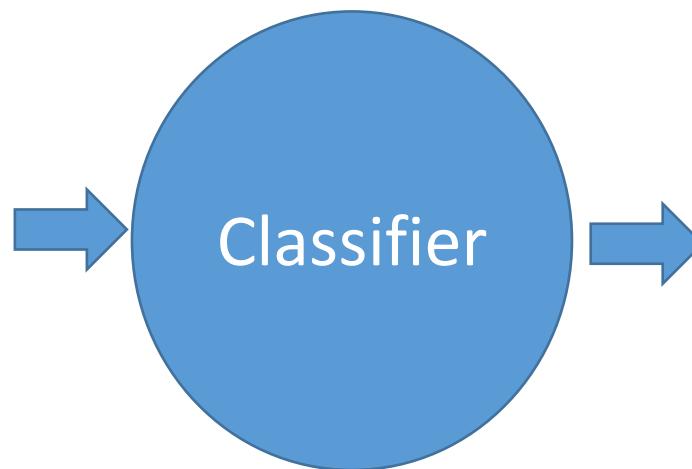
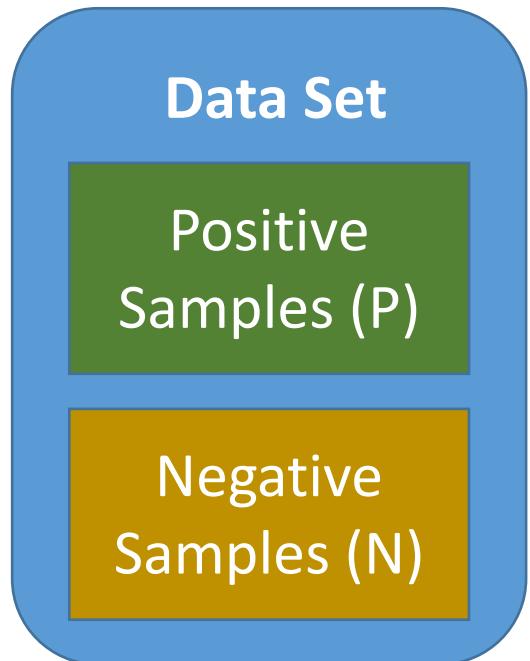
Objective Space



- The Evaluation of a Genome associates a genome/individual (set of parameters for GA or string for GP) with a set of scores
- What are these scores
 - True Positive – TP
 - How often are we identifying the desired object
 - False Positive – FP
 - How often are we identifying something else as the desired object
 - More? – On board, next page
- Objectives
 - Set of measurements each genome (or individual) is scored against
 - Phenotype
- Objective Space – Set of objectives
- Evaluation – Maps an genome/individual
 - From a location in search space
 - Genotypic description
 - To a location in objective space
 - Phenotype description

Classification Measures

Confusion Matrix



	Predicted: Positive	Predicted: Negative
Actual Positive (P)	True Positive (TP)	False Negative (FN)
Actual Negative (N)	False Positive (FP)	True Negative (TN)

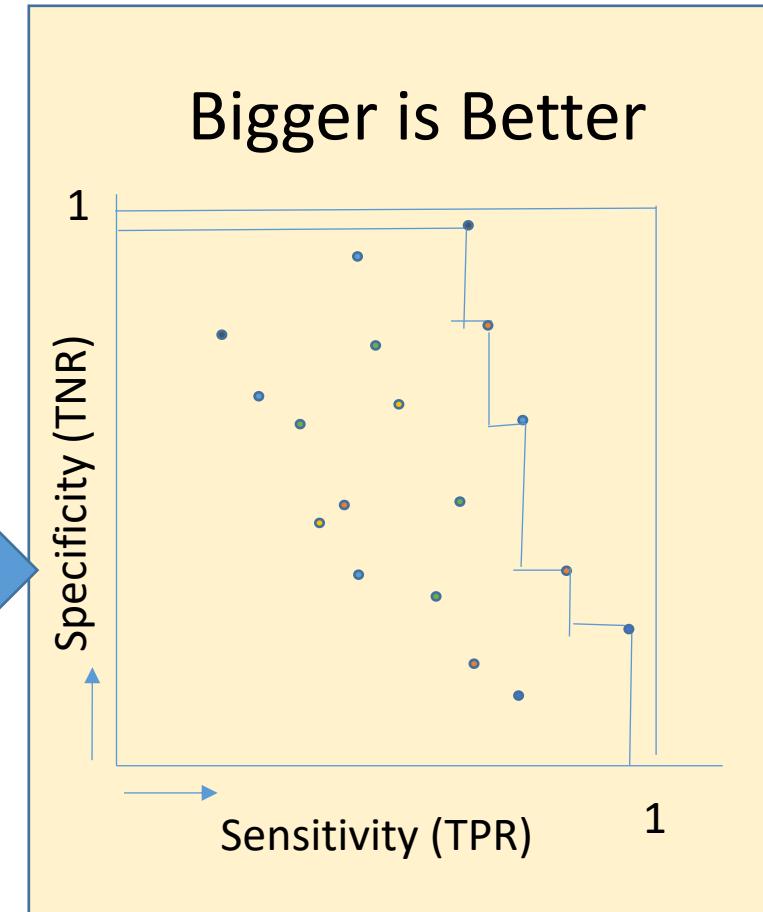
Board Example

Maximization Measures

	Predicted: Positive	Predicted: Negative
Actual Positive (P)	True Positive (TP)	False Negative (FN)
Actual Negative (N)	False Positive (FP)	True Negative (TN)

- Sensitivity or True Positive Rate (TPR)
 - AKA hit rate or recall
 - $TPR = TP/P = TP/(TP+FN)$

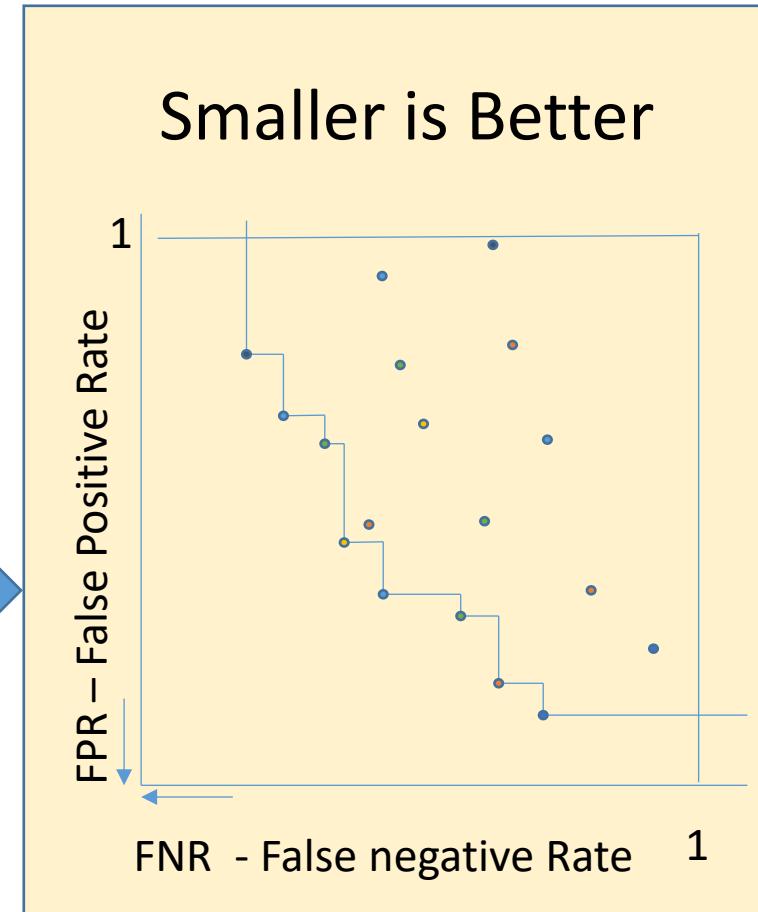
- Specificity (SPC) or True Negative Rate (TNR)
 - $TNR = TN/N = TN/(TN+FP)$



Minimization Measures

	Predicted: Positive	Predicted: Negative
Actual Positive (P)	True Positive (TP)	False Negative (FN)
Actual Negative (N)	False Positive (FP)	True Negative (TN)

- False Negative Rate (FNR)
 - $FNR = FN/P = FN/(TP+FN)$
 - $FNR = 1 - TPR$
- Fallout or False Positive Rate (FPR)
 - $FPR = FP/N = FP/(FP+TN)$
 - $FPR = 1 - TNR = 1 - SPC$



Other Measures

	Predicted: Positive	Predicted: Negative
Actual Positive (P)	True Positive (TP)	False Negative (FN)
Actual Negative (N)	False Positive (FP)	True Negative (TN)

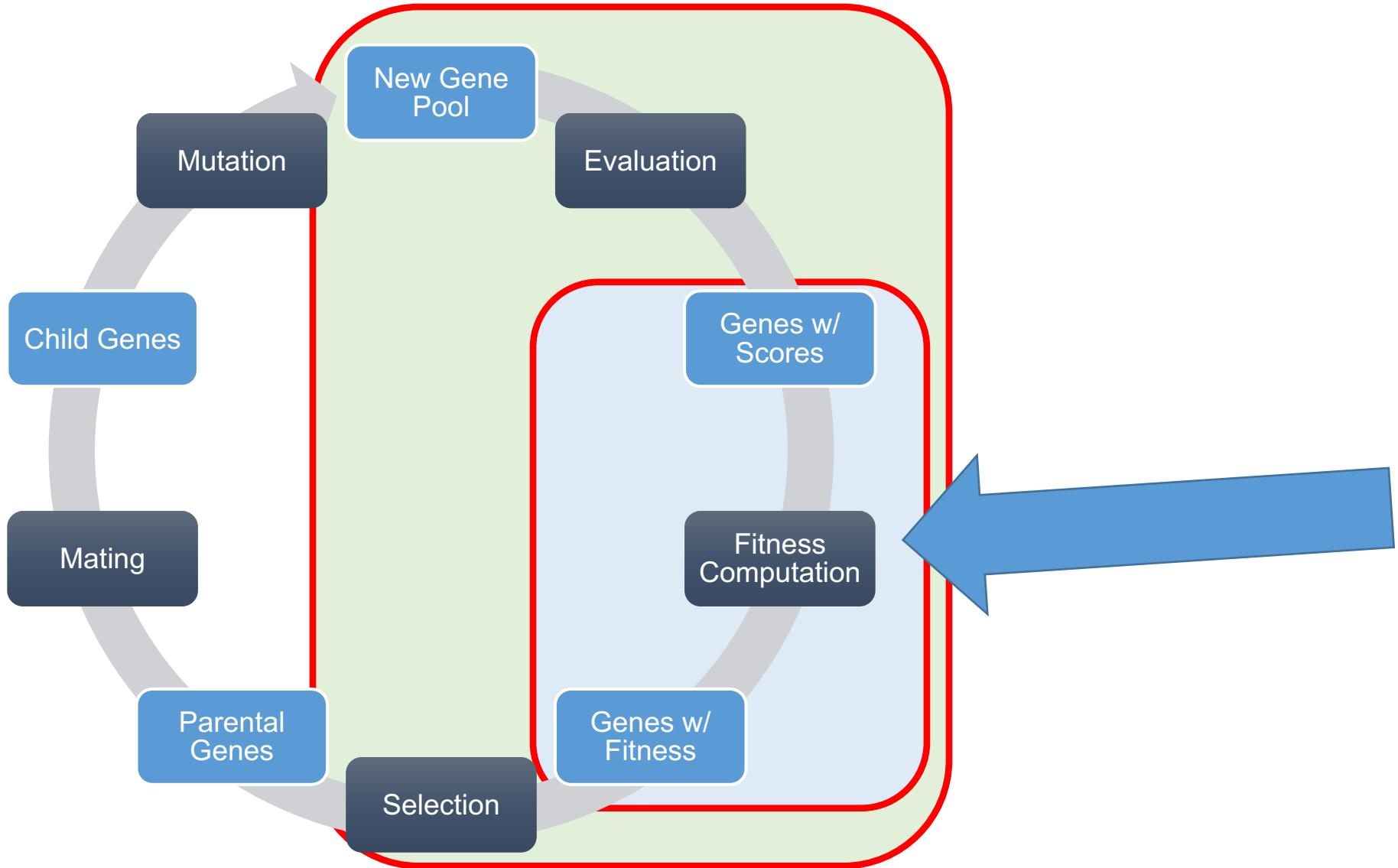
- Precision or Positive Predictive Value (PPV)
 - $PPV = TP / (TP + FP)$
 - Bigger is better
- False Discovery Rate (FDR)
 - $FDR = FP / (TP + FP)$
 - $FDR = 1 - PPV$
 - Smaller is better

- Negative Predictive Value (NPV)
 - $NPV = TN / (TN + FN)$
 - Bigger is better

- Accuracy (ACC)
 - $ACC = (TP+TN) / (P+N)$
 - $ACC = (TP+TN) / (TP + FP + FN + TN)$
 - Bigger is better

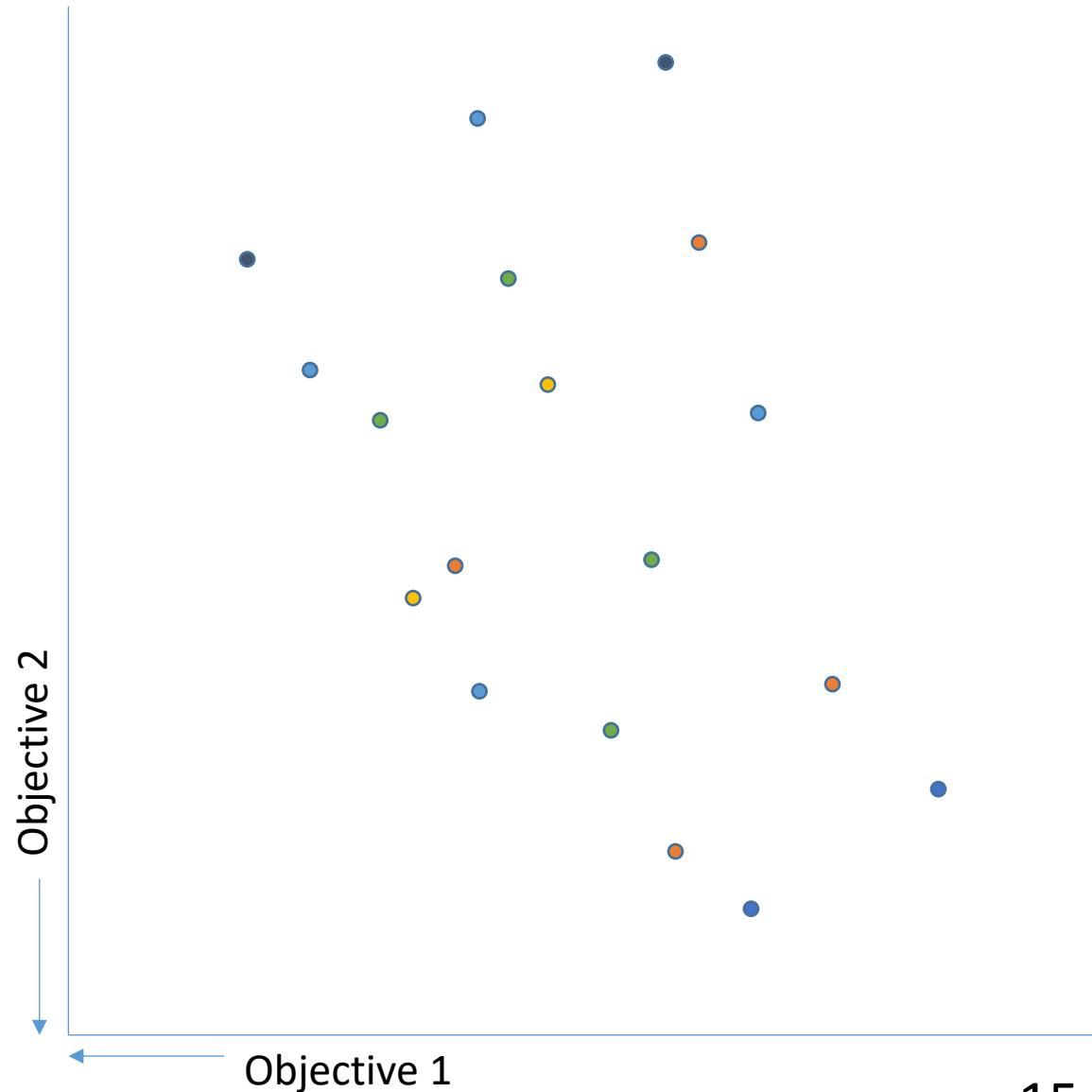
Example from Class Population: Python Versus Machine Learning

Lecture Scope



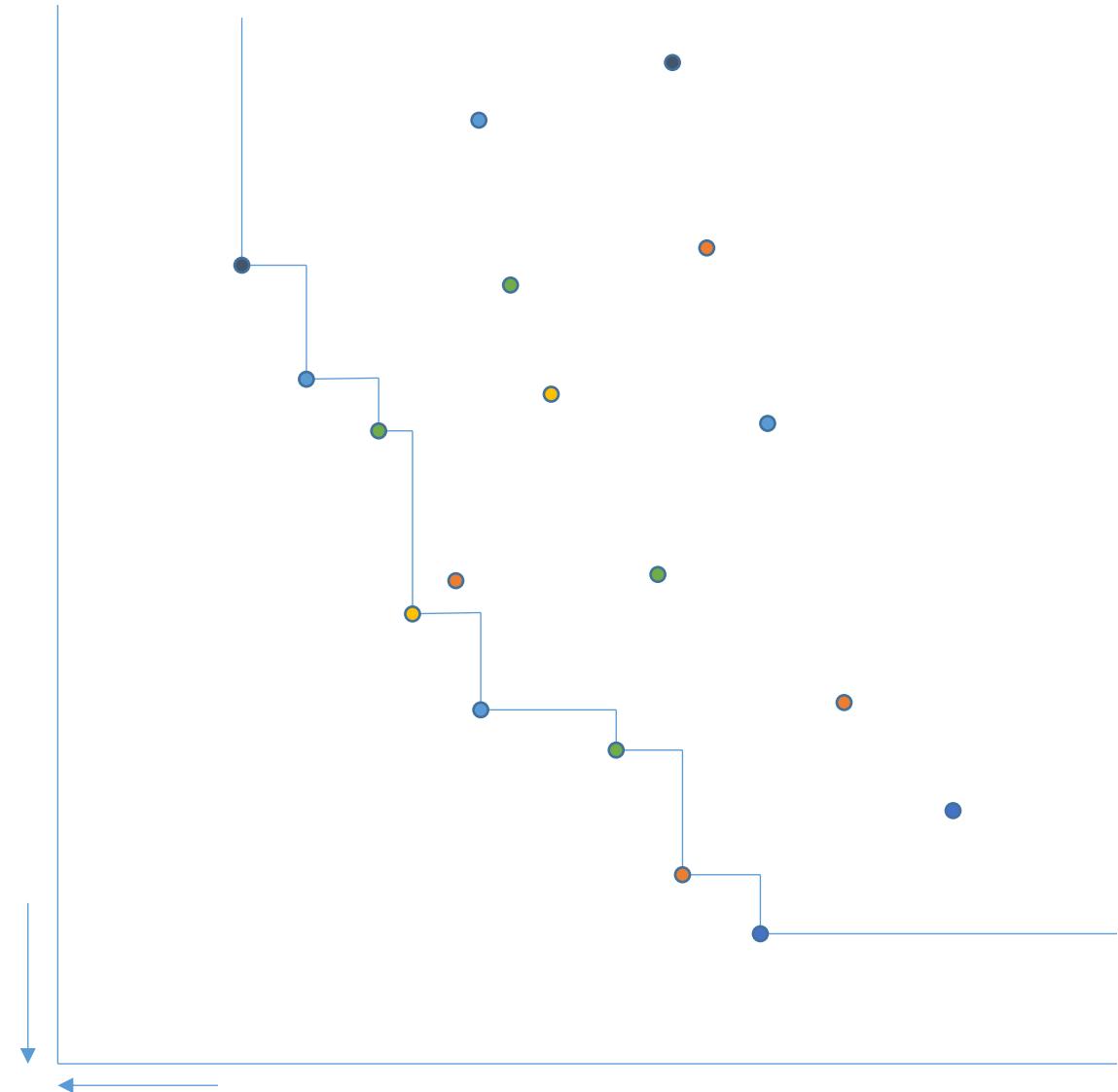
Objective Space

- Each individual is evaluated using objective functions
 - Mean squared error
 - Cost
 - Complexity
 - True positive rate
 - False positive rate
 - Etc...
- Objective scores give each individual a point in objective space
- This may be referred to as the **phenotype** of the individual
- Examples are shown with two objectives, but all techniques we will discuss are extensible to N objectives



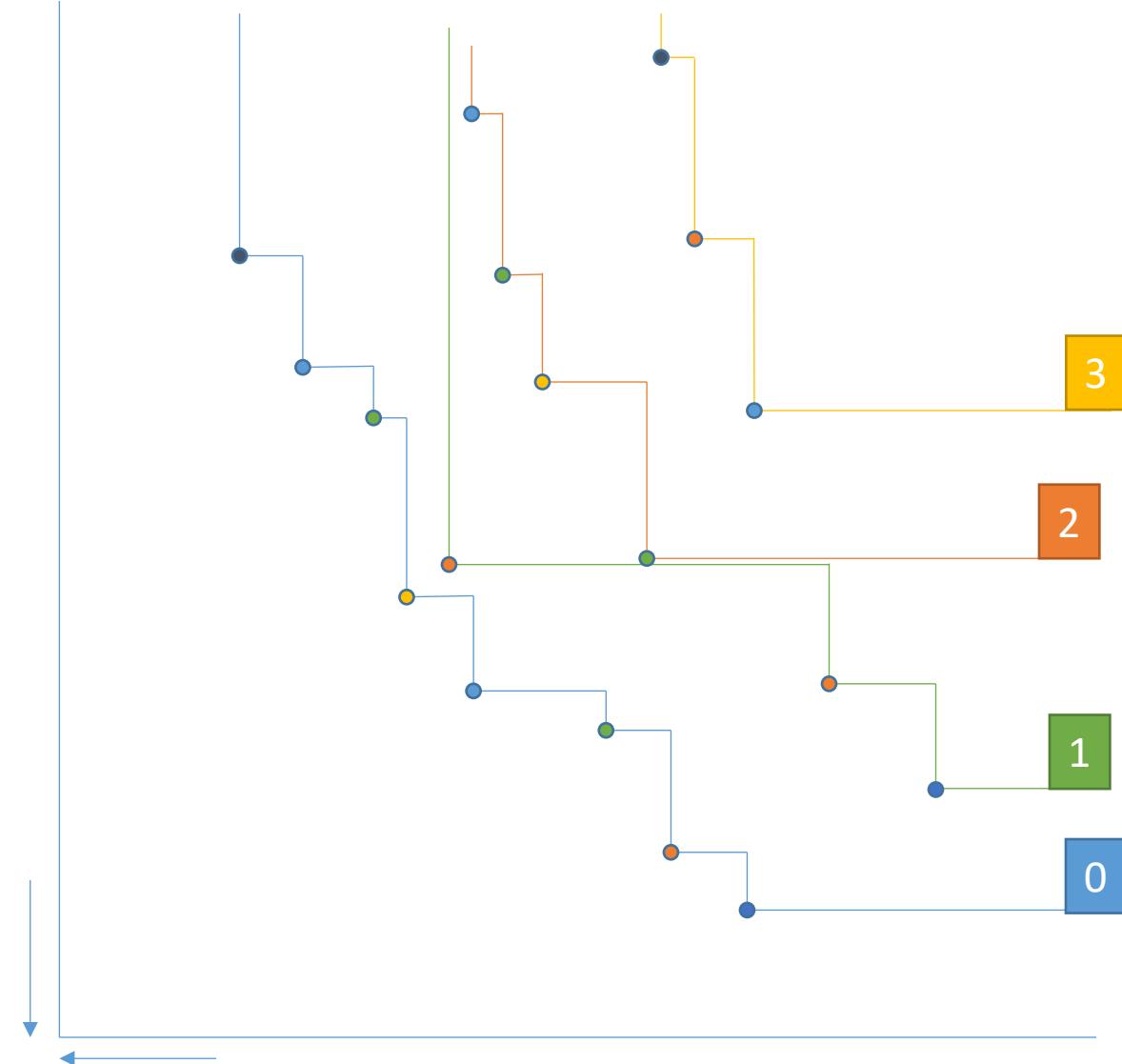
Pareto Optimality

- An individual is **Pareto** if there is no other individual in the population that outperforms the individual on all objectives
- The set of all Pareto individuals is known as the **Pareto frontier**
- These individuals represent unique contributions
- We want to drive selection by favoring Pareto individuals
 - But maintain diversity by giving all individuals some probability of mating



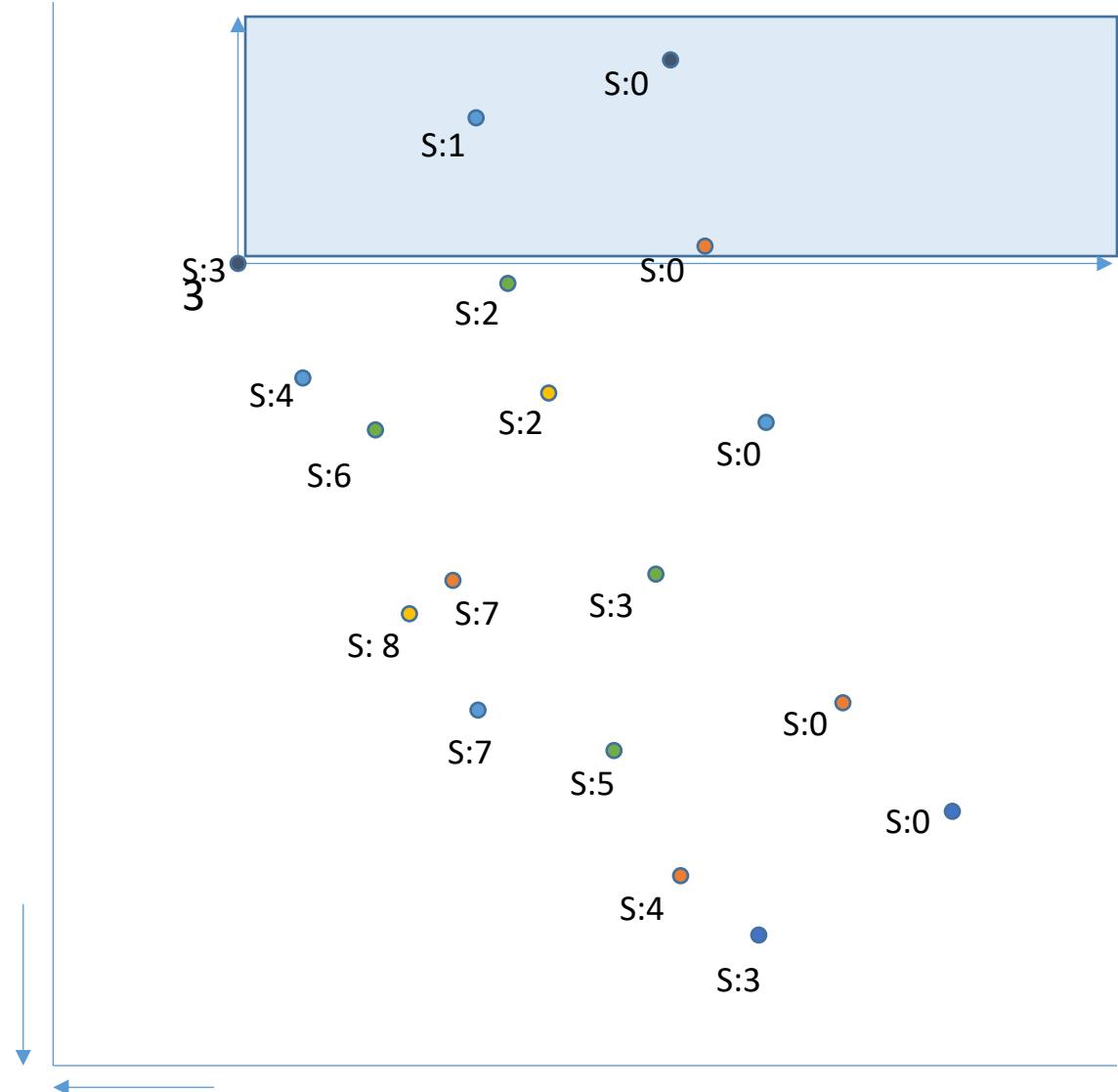
Nondominated Sorting Genetic Algorithm II (NSGA II)

- Population is separated into nondomination ranks
- Individuals are selected using a binary tournament
- Lower Pareto ranks beat higher Pareto ranks
 - For example, an individual on the blue front will beat out an individual on the orange front
- Ties on the same front are broken by **crowding distance**
 - Summation of normalized Euclidian distances to all points within the front
 - Higher crowding distance wins



Strength Pareto Evolutionary Algorithm 2 (SPEA2)

- Each individual is given a strength S
 - S is how many others in the population it dominates



Strength Pareto Evolutionary Algorithm 2 (SPEA2)

- Each individual is given a strength S
 - S is how many others in the population it dominates
- Each individual receives a rank R
 - R is the sum of S 's of the individuals that dominate it
 - Pareto individuals are nondominated and receive an R of 0
- A distance to the k^{th} nearest neighbor (σ^k) is calculated and a fitness of $R + 1/(\sigma^k + 2)$ is obtained

