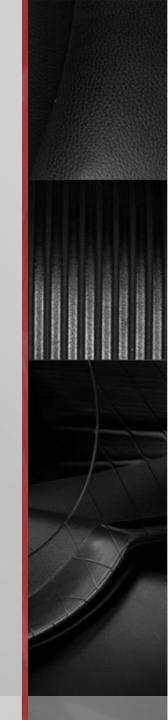
## CSCI 446: Artificial Intelligence

**Genetic Algorithms** 



### Basic concept is Darwin's theory of evolution

- Evolution is the process of a species adapting to and surviving in a given environment
- The more able to survive (the fittest) are also the more likely to reproduce, thus improving the gene pool over time

- Rattlesnake example
  - Some rattlesnakes don't have rattles
  - In AZ? NM? This percentage has been increasing.
  - The theory is that if people hear the rattle, the snake is more likely to get killed, so snakes without rattles are more likely to mate and reproduce and their offspring are more likely not to have rattles

- Two main issues of simulating evolutionary behavior
  - Encoding how do we represent a problem as a chromosome containing a string of genes
  - Evaluation how do we tell which individuals (chromosomes) are more fit than others, and therefore allowed to reproduce

- Step 1:
  - Represent the domain as a chromosome
  - Choose population size N
  - Choose crossover probability Pc
  - Choose mutation probability Pm

- Step 2:
  - Define fitness function to evaluate the "goodness" of a particular solution
- Step 3:
  - Randomly generate initial population of N chromosomes
- Step 4:
  - Calculate the fitness of each chromosome

- Step 5:
  - Select a pair of chromosomes for mating

Base selection on both their fitness and a probability function

- Step 6:
  - Create offspring by applying crossover and mutation operators
- Step 7:
  - Put new offspring in the population
- Step 8:
  - Repeat. Go back to step 5 until new population is of size N

- Step 9:
  - Replace parent population with child population
- Step 10:
  - Go to step 4 and repeat process until termination criterion is reached

# Comparison to Traditional Optimization and Search

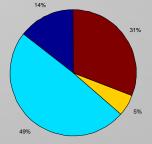
- GAs work with an encoding of the parameter set, not the parameters themselves
- GAs search from a population of points, not a single point
- GAs use a payoff / objective function, not derivatives of other auxiliary knowledge
- GAs use probabilistic transition rules, not deterministic rules

- Selection / Reproduction
  - Cloning a chromosome to put in the mating pool
  - Recall that one of the main issues in evolutionary computing is how to represent the problem parameters as a chromosome / the eventual solution
  - The second main issue is how to evaluate the goodness, or fitness, of that chromosome / solution
  - When you have these two things, you can then see how operators work
  - In reproduction, each chromosome is evaluated according to a fitness function
  - The value of its fitness influences its probability of being chosen for reproduction, that is, carrying its genes to the next generation

### Example

Number	String	Fitness	% of Total
1	01101	169	14.4
2	11000	576	49.2
3	01000	64	5.5
4	10011	361	30.9
Total		1170	100.0

- 2<sup>nd</sup> one has a greater chance of being chosen for reproduction
- Can visualize this as a roulette wheel selection



#### Crossover

Proceeds in two steps

Members of the mating pool are mated at random

Each pair of strings / chromosomes is crossed

Choose an integer position k

Uniform choice between 1 and  $\ell$ -1 where  $\ell$  is the length of the chromosome

Create two more strings by swapping all genes / characters from k+1 to  $\ell$  inclusively

e.g. choose k=4 (position 4)

A1 = 01101 A2 = 11000 A1' = 01100 A2' = 11001

### Mutation

- Actually plays a minor role
- Needed because sometimes good genetic material is lost, but most of the time, mutation has a degrading effect on the fitness
- Mutation is the occasional (small probability) random alteration of the value of a string position
- Frequency of mutation usually on the order of 1 per 1000 bit transfers
  Actually even smaller probability in natural populations

- Where is this used?
  - Nathan Fortier's Open Song Composer