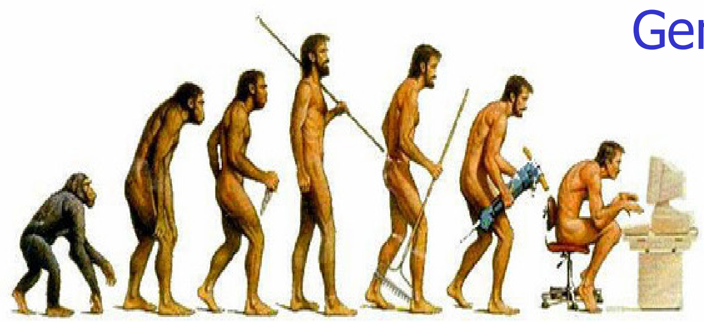
CSE4403 3.0/CSE6602E - Soft Computing
Winter 2011



Lecture 9

Genetic Algorithms & Evolution Strategies

Guest lecturer: Xiangdong An
xan@cs.yorku.ca

Genetic algorithms

- J. Holland, Adaptation in Natural and Artificial Systems, 1975.

Sketch of the simple GA

Representation	Bit-strings
Recombination	1-point crossover
Mutation	Bit flip
Parent selection	Fitness proportional
Survival selection	Generational

Simple genetic algorithms – an example

- Maximizing x^2 for x in the range 0-31, represented as a 5-bit binary string

genotype — phenotype — $Prob_i = f_i / \Sigma f_i$ — $f_i / \bar{f}$

String no.	Initial population	x value	Fitness $f(x)=x^2$	$Prob_i$	Expected count	Actual count
1	01101	13	169	0.14	0.58	1
2	11000	24	576	0.49	1.97	2
3	01000	8	64	0.06	0.22	0
4	10011	19	361	0.31	1.23	1
Sum			1170	1.00	4.00	4
Average			293	0.25	1.00	1
Max			576	0.49	1.97	2

Initialization, and parent selection

Simple genetic algorithms – an example

String no.	Mating pool	Crossover point	Offspring after xover	*x* value	Fitness $f(x) = x^2$
1	0110|1	4	01100	12	144
2	1100|0	4	11001	25	625
2	11|000	2	11011	27	729
4	10|011	2	10000	16	256
Sum					1754
Average					439
Max					729

Crossover and offspring evaluation

Simple genetic algorithms – an example

String no.	Offspring after xover	Offspring after mutation	*x* value	Fitness $f(x)=x^2$
1	01100	11100	28	784
2	11001	11001	25	625
2	11011	11011	27	729
4	10000	10100	18	324
Sum				2462
Average				615.5
Max				784

Mutation and offspring evaluation

Representation of individuals

- Binary representations
  - Gray coding ensures consecutive integers always have hamming distance one, i.e., two successive values differ in only one bit
  - 00,01,11,10
  - 000,001,011,010,110,111,101,100
- Integer representations
- Real-valued of floating point representations
- Permutation representations
  - Eight-queens example

Mutation

- For binary representation

1	0	1	0	0	0	0	1	0	1

⇩

1	0	0	1	0	0	0	0	0	1

Bitwise mutation rate p_m

Given sequence length L, $L \cdot p_m$ bits will be mutated averagely

Mutation

- For integer representation
  - Random setting (cardinal)

3	2	9	8	0	1	4	5	7	9

⇩

3	2	8	5	0	1	4	6	7	9

Bitwise mutation rate p_m

Given sequence length L, $L \cdot p_m$ bits will be mutated averagely

  - Creep mutation (ordinal)
    - Step sizes

Mutation

- For floating-point representation
  - $<x_1, \ldots, x_n> \Rightarrow <x_1', \ldots, x_n'>$ $\quad x_i, x_i' \in [L_i, U_i]$

Uniform mutation

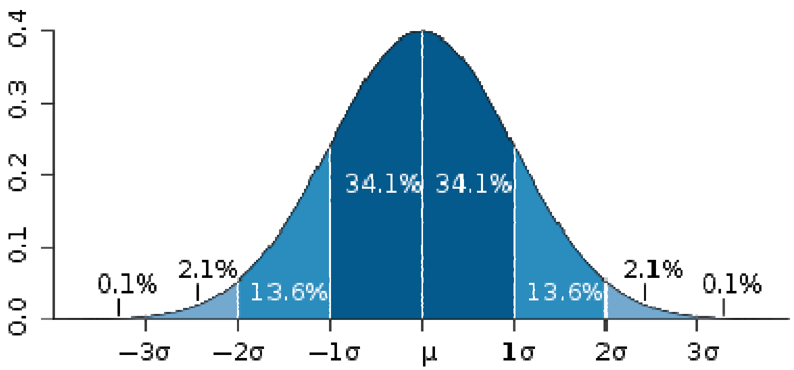
Uniformly form range $[L_i, U_i]$

Nonuniform mutation with a fixed distribution

Gaussian distribution

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

$N(\mu,\sigma^2)$



Cauchy distribution

Mutation

- For permutation representation
  - Swap mutation

1	2	3	4	5	6	7	8	9
1	5	3	4	2	6	7	8	9

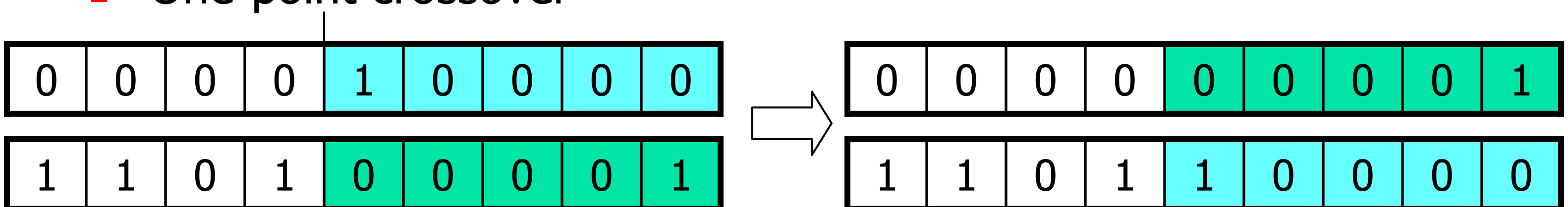
  - Insert mutation

1	2	3	4	5	6	7	8	9
1	2	5	3	4	6	7	8	9

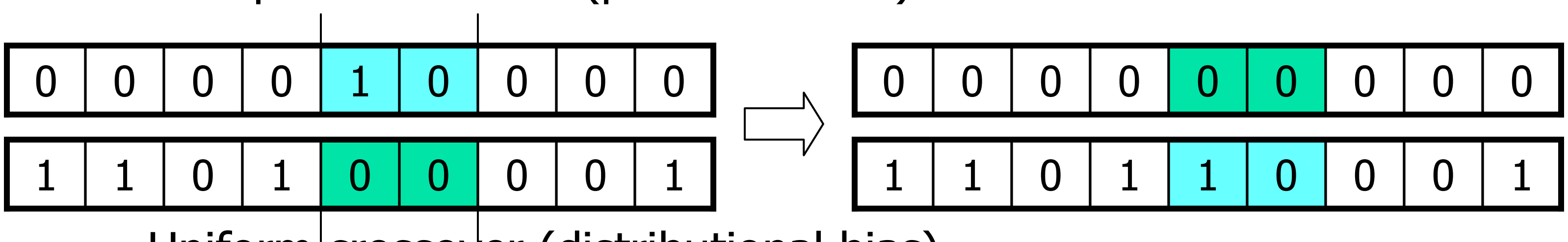
  - Scramble mutation
    - Have positions of a subset values scrambled
  - Inversion mutation
    - Reversing the order of a range of values in the string

Recombination

- Crossover rate p_c
  - If lower than p_c, recombination happens; otherwise asexually
- For binary (integer, floating-point) representation
  - One-point crossover



  - Two-point crossover (positional bias)



  - Uniform crossover (distributional bias)
    - For each gene position, a random number generated from [0,1).
    - If smaller than, say 0.5, inherit from parent1, otherwise parent2.

Recombination

- For floating-point representation
  - $z_i = ax_i + (1-a)y_i$, $a \in [0, 1]$
  - Simple recombination
    - Partial genes generated by arithmetic recombination
  - Single arithmetic recombination
    - Only one gene generated by arithmetic recombination
  - Whole arithmetic recombination
    - All genes are generated by arithmetic recombination

Range of gene values in the population
will be reduced due to averaging

Population models

- The generation model
  - After each generation, the whole population is replaced by its offspring
- The steady-state model
  - After each generation, only partial population is replace by its offspring

Parent selection

- Fitness proportional selection
  - Premature convergence
    - Outstanding individuals take over the entire population very quickly, i.e., genetic drift
    - No selection pressure at later stage
- Ranking selection
  - Allocating selection probability based on ranking of individuals, not fitness values
    - Preserves a constant selection pressure

Survivor selection

- Age-based replacement
- Fitness-based replacement

Evolution strategies

- By Rechenberg and Schwefel, 1960s

Sketch of ES

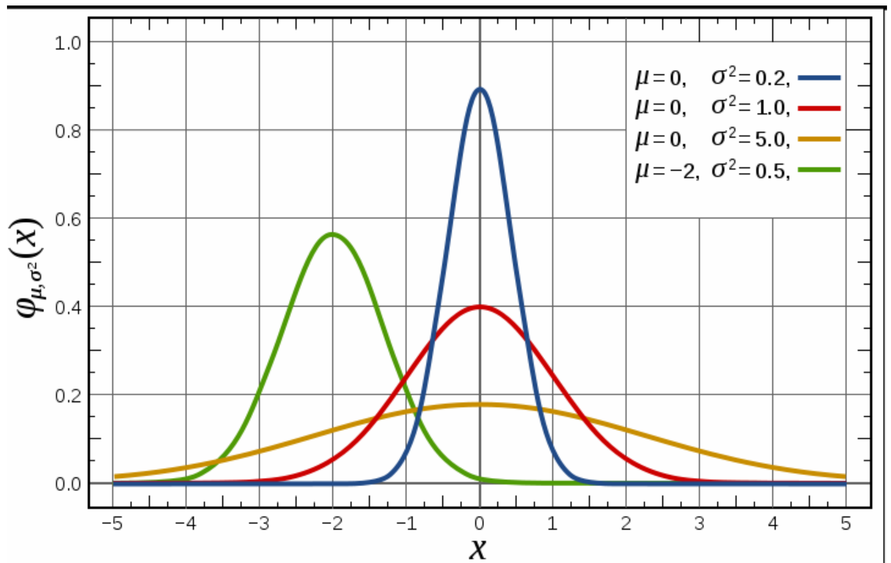
Representation	Real-valued vectors
Recombination	Discrete or intermediatry
Mutation	Gaussian perturbation
Parent selection	Uniform random
Survival selection	(μ,λ) or $(\mu+\lambda)$
Speciality	Self-adaptation of mutation step sizes

Representation

- Typically used for continuous parameter optimization
  - A vector of floating-point variables $<x_1, \ldots, x_n>$
  - Objective function: $R^n \rightarrow R$
- To self-adapt
  - $<\underbrace{x_1, \ldots, x_n}_{x}, \underbrace{\sigma_1, \ldots, \sigma_m}_{\sigma}, \underbrace{\alpha_1, \ldots, \alpha_m}_{\alpha}>$, $1 \leq m \leq n$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

σ determines the perturbation extent, often called mutation step size.



Mutation

$x_i' = x_i + N(0, \sigma_i')$

Evolve σ_i to get σ_i' first, and then use it to mutate x_i

Self-adaptation adjusts the mutation strategy as the search is proceeding. The underlying assumption is that at different stages of evolution, different mutation may be appropriate.

Uncorrelated mutation with one step size

- The same distribution is used to mutate each x_i, therefore only one σ for each individual

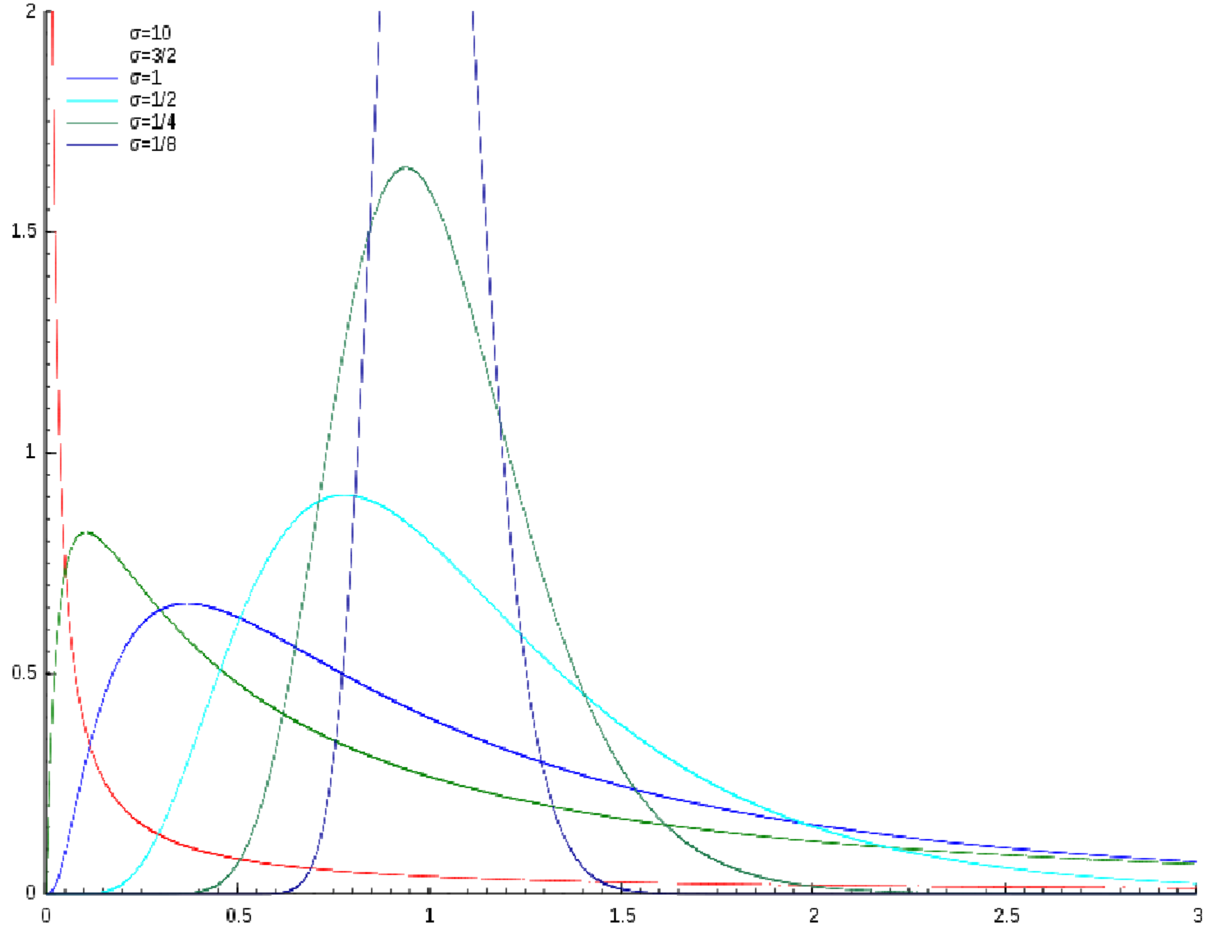
σ is mutated by a term e^{Γ}, with Γ a random variable drawn each time from a normal distribution with mean 0 and standard deviation τ

$\Gamma \sim N(0, \tau)$

e^{Γ} is a lognormal distribution

$\sigma' = \sigma \cdot e^{\Gamma}$

Lognormal distribution



Recombination

- Intermediary recombination
  - $z_i = ax_i + (1-a)y_i$,
- Discrete recombination
  - $z_i = x_i$ or y_i, chosen randomly

Survivor selection

- After creating λ offspring, the best μ of them are chosen deterministically:
  - From the offspring only, called (μ, λ) selection
    - Discards all parents and in principle is able to leave local optima
  - From the union of parents and offspring, called $(\mu+\lambda)$ selection
    - Preserves outdate solutions
    - Misadapted parameters may survive to relatively larger number of generations when having good object variables but bad strategy parameters