Lecture Series on

Intelligent Control

Lecture 24
Evolutionary Strategies and Evolutionary Programming

Kwang Y. Lee
Professor of Electrical & Computer Engineering
Baylor University
Waco, TX 76798, USA
Kwang_Y_Lee@baylor.edu

1

1

Evolutionary Strategies and Evolutionary Programming

A bit of history

- EVOLUTION STRATEGIES
  - Rechenberg, Schwefel,
  - Technical University of Berlin, 1963
  - https://en.wikipedia.org/wiki/Ingo_Rechenberg

- EVOLUTIONARY PROGRAMMING
  - L. Fogel,
  - "Artificial Intelligence Through Simulated Evolution
  - https://en.wikipedia.org/wiki/Lawrence_J._Fogel
- Are they really distinct?

2

2

INDIVIDUALS AND POPULATIONS

- An individual stands for an alternative or a solution
- An individual is represented by a real-valued vector

  (but in modern models it can also be an integer-valued or mixed integer-valued vector)
- A population is a set of individuals considered belonging to a same generation
- Offsprings are individuals (alternatives, solutions) derived from other solutions by means of *mutation* and *recombination*
- Fitness is a measure of adaptation to an environment or landscape - defined by the objective function in classical optimization

3

3

ES GENERAL PROCESS

- Generate initial population of μ individuals
- Evaluate fitness of the existing μ individuals

- DO
- Generate, from existing μ individuals, λ offspring (use recombination, then mutation)
- Evaluate fitness of new individuals
- Select μ survivors to form the next generation
- LOOP *until termination criterion is met*

4

4

(μ *and* λ) ES
EVOLUTION STRATEGIES

- (μ,λ) ES - an evolutionary process where a population with μ individuals (parents) originates λ descendents (children), from which μ are selected to form a new population (a new generation).

- (μ+λ) ES - an evolutionary process where a population with μ individuals (parents) originates λ descendents (children); and from the set of {μ+λ} individuals, μ are selected to form a new population (a new generation).

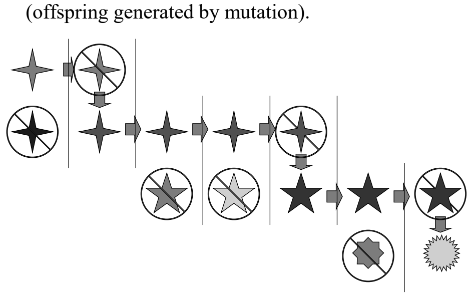
- Children are created by mutation and recombination.

5

5

THE FIRST: (1+1) ES

- The first historical model was a (1+1)ES: 1 parent, 1 offspring, fitness calculated for both, the best one selected (offspring generated by mutation).



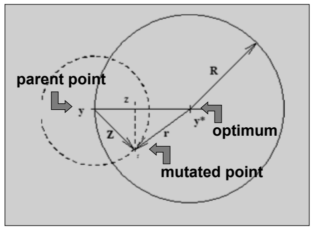
6

6

The spherical model

$$F_1(\mathbf{y}) = \sum_{i=1}^{n} \left(y_i - y_i^*\right)^2$$

- An isotropic fitness landscape
- Mutated point $\tilde{y} = y + Z$
- Z is a random vector
- R is the radius of the domain of success of mutations
- Rate of progress
  $\varphi = E\{R-r\}$



7

7

THE FIRST: (1+1) ES

- The selection scheme of the (1+1)ES is elitist.
  The mutation scheme is based on Gaussian mutations.
- A solution X is a vector with *n* real components
- A mutation is obtained by adding a random vector Z

x_1	x_2	...	x_n

$$\mathbf{Z} = \sigma\left(N_1(0,1),\ldots,N_n(0,1)\right)^t$$

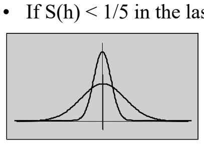
- Each component is mutated according to a Gaussian distribution with 0 mean and variance 1
- σ is the *step size* or *mutation strength*

8

8

The 1/5 Rule

- To reach the optimum, the mutation strength σ cannot remain constant
  - The 1/5 Rule (Rechenberg): the Success Rate S(h) (ratio between successful mutations and all mutations) must be close to 1/5
- If S(h) > 1/5 in the last h generations ⇒ increase σ
- If S(h) < 1/5 in the last h generations ⇒ decrease σ



- Optimal value for h: h = 10n
- (n - dimension of the search space)

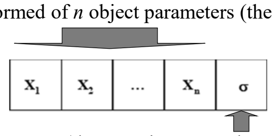
$$\mathbf{Z} = \sigma\left(N_1(0,1),\ldots,N_n(0,1)\right)^t$$

9

9

The (1,λ) ES

- In the (1,λ) ES, 1 individual mutates to generate λ offspring and only 1 of these survives to the next generation. (parent and children do not compete)
- Each individual is formed of *n* object parameters (the variables)



- and one strategic parameter (the mutation strength, equal for all individuals)

10

10

The σSA (1,λ) ES

- σSA, a self-adapting scheme: the mutation strength evolves and is subject to selection
- The object parameters are added with Gaussian perturbations, but...
- The strategic parameter is subject to a multiplicative mutation

$$\sigma_k^{(g+1)} = \xi\sigma^{(g)}$$

$$\xi = e^{\tau N(0,1)}$$

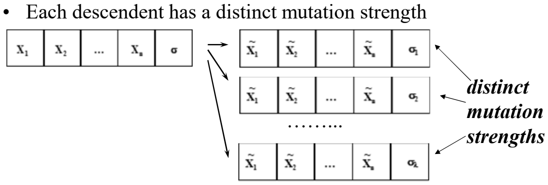
- *ξ is a random number with E{ξ}=1*
- *ξ may follow a lognormal distribution*
- *or a symmetrical two-point distribution*

11

11

The σSA (1,λ) ES

- Each descendent has a distinct mutation strength



- The selection process also selects the most favorable mutation strength...

12

12

The σSA (1,λ) ES

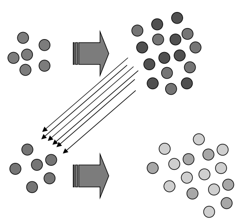
- How to select the learning parameter τ? $\xi = e^{\tau N(0,1)}$
  - Schwefel's rule: $\tau \sim (n)^{-1/2}$, with n being the dimension of the search space
  - A model like this displays a linear convergence order
    - with nearly optimum convergence rate

13

13

The σSA (μ,λ) ES

- The (μ,λ) ES is a natural extension of the (1,λ) ES
- One may select individuals randomly to reproduce, or
- one may give preference to the most fit, by:
  - tournament selection
  - roulette selection

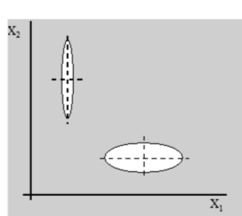


14

14

Search patterns in σSA (μ,λ) ES

- Instead of a single strategic parameter, one may have a mutation strength associated with each variable
- Mutations in the variables of an individual are uncorrelated: they adapt to an anisotropic fitness landscape, but search proceeds along coordinate axes



15

15

$(\mu+\lambda)$ ES: Evolutionary Programming

- Make $\mu = \lambda$ - this ES becomes EP!
- Major historical difference: German researchers gave preference to elitist selection, while Fogel followers adopted Stochastic Tournament (inspired by GA?)

- Simple Stochastic Tournament:
  - *Randomly select two individuals*
  - *Compare fitness*
  - *Select to survive (to be reproduced) the best fit, with a given high probability*
  - *Repeat until λ individuals are selected*

16

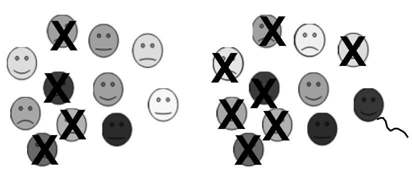
16

Evolutionary Programming

- Generate Initial Population

Replicate the Population and Mutate the offspring

Compute fitness and Select Best Individuals



17

17

EP and self adaptation

- Self-adaptation schemes have been adopted in EP.
- However, instead of Lognormal mutations, EP has traditionally adopted Gaussian mutations, still multiplicative

$$\sigma^{(g+1)} = \xi\sigma^{(g)} \qquad \xi = 1 + \tau N(0,1)$$

- But…
- it has been shown that this EP scheme is an approximation of the general ES scheme using the Lognormal distribution

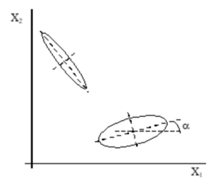
18

18

Enhancing the mutation scheme

- New strategic parameters:
- correlation between mutations

- this allows a more robust search
- pattern in complex landscapes



$$\widetilde{\mathbf{X}} = \mathbf{X} + \mathbf{CZ}$$

- Z is a random vector and C is a Matrix defining a Mahalanobis metric in space - it may be defined by elementary angles α, subject to mutations
- https://en.wikipedia.org/wiki/Mahalanobis_distance

19

19

Recombination: $(\mu/\rho,\lambda)$ ES

- Recombination: mixing ρ individuals (among μ parents) to obtain a new one
- Biology is based on $\rho = 2$ but… why stop at 2?

- Recombination variants:
  - uniform crossover
  - intermediary recombination
    - global
    - local
  - point crossover

20

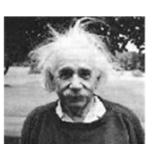
20

The importance of recombination

Max the success in life f
$f(x,y) = x^2 + y^2$ $x,y \in [0,100]$

45	**38**

X (Intelligence) **Y (Beauty)**



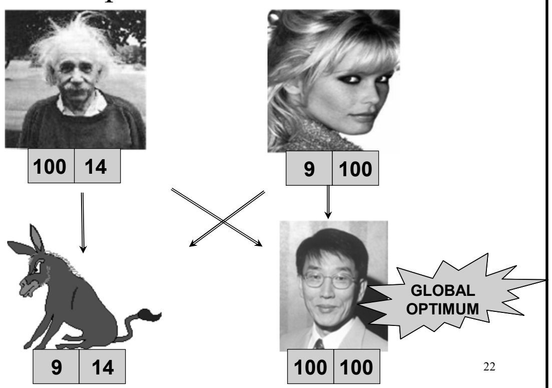
100	**14**

9	**100**

21

21

The importance of recombination



22

22

Recombination variants

- Uniform crossover: a new individual is formed by randomly selecting each component from a different parent.

15	28	-7	…	45
33	-2	33	…	78
43	12	-8	…	73
43	-2	-8	…	45

- Example for $\rho = 3$

23

23

Recombination variants

- Intermediary recombination: each variable in the new individual is a combination of the values of the same variable in the ρ parents
  - *there are many variants of intermediary recombination*
- Example for $\rho = 2$: for the variable k,

$$x_k^{new} = u_k x_{k,j1} + (1 - u_k) x_{k,j2}$$

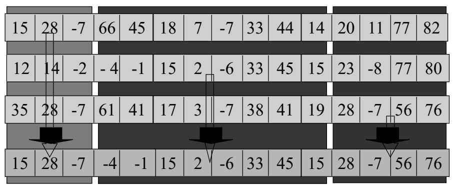
- where u_k is a random number sampled from a uniform distribution in [0,1]

24

24

Recombination variants

- Point crossover: the parents are sliced in parts (defined by randomly selecting γ points, common to all parents) and the new individual receives a part from each parent, in turns
  - *this is what happens in Genetic Algorithms, for ρ = 2*



25

25

Handling constraints

- As in all evolutionary algorithms…
- By controlling mutations so that no unfeasible individuals are generated
- By applying penalties to the fitness value of unfeasible individuals

26

26

Starting point

- Typically:
- By randomly generating the coordinates of the μ first individuals
- By generation mutations from a seed or initial individual

27

27

Conclusions

- Evolution Strategies, as a general algorithmic approach, is very successful
- Among Evolutionary Algorithms, it ranks in a high place - there are many reports and examples where an ES model has performed better than a Genetic Algorithm model
- There is, however, no demonstration of definite superiority - ultimately, it depends in a high degree of the skill and cleverness of the engineer or the scientist, when developing the model

28

28