Evolutionary Algorithms

Homework

Second task

Author
Carlos Sánchez Páez
http://www.github.com/csp98

Budapest University of Technology and Economics

Academic year 2018-2019

1. Let us suppose, that an algorithm's running time is polynomial, that is cn^α for some α, $c \in \mathbb{R}$ constants. Give an estimate for c and α if for input lengths n = [4, 5, 6, 7, 8, 9, 10] we measured the following running times [37.1 58.7 84.0 115.1 150.8 190.9 235.2].

   We will solve the exercise with the help of GNUPLOT. First, we will define the function and then tell the program to fit it to the given values:

```
f(x) = c * x ** a
fit f(x) 'ex1_data.txt' via a,c
```

   The final set of parameters obtained is the following:

	a	c
Value	2.00895	2.3073
Error	0.25%	1.088%

   So our function is $f(n) = 2.3073n^{2.00895}$.

2. Let us suppose, that we have a population containing 4 individuals called e_1, e_2, e_3, e_4. Their fitness's are 0.4, 0.7, 0.3, 0.05. We use a roulette-wheel selection to select the four parents. Answer these two questions for $f(e)$ and also for the scaled $\hat{f}(e) = f^2(e)$ fitness function.

	$f(e_i)$	$f^2(e_i)$
e_1	0.4	0.16
e_2	0.7	0.49
e_3	0.3	0.09
e_4	0.05	0.0025
Sum	1.45	0.7425

   - What is the probability, that e_2 won't be chosen as parent at all?

     To solve this question we will use the complement rule of probability. If $P(e_2)$ is the probability of electing e_2, $P(\overline{e_2}) = 1 - P(e_2)$ is the probabilty of not choosing it at all. We will start calculating $P(e_2)$.

$$P(e_2) = \frac{f(e_2)}{\sum_{i=1}^{n} f(e_i)} = \frac{0.7}{1.45} = 0.482758621$$

$$P(\overline{e_2}) = 1 - P(e_2) = 1 - 0.482758621 = 0.517241379$$

     So the probability is **51.72%**.

     For $f^2(e)$ we apply the same procedure:

$$P'(e_2) = \frac{f^2(e_2)}{\sum_{i=1}^{n} f^2(e_i)} = \frac{0.49}{0.7425} = 0.65993266$$

$$P'(\overline{e_2}) = 1 - P'(e_2) = 1 - 0.65993266 = 0.34006734$$

     So the probability is **34.01%**.

- What is the probability, that e_3 will be chosen two times?

  The probability of the intersection is defined by the following statement: $P(A \cap B) = P(A) \cdot P(B/A)$. $P(B/A)$ means the probability of B knowing that A happened. In this case, the selection of the parents is made in an *equiprobable* way, so $P(B/A) = P(B)$.

$$P(e_3 \cap e_3) = (P(e_3))^2 = (\frac{f(e_3)}{\sum_{i=1}^{n} f(e_i)})^2 = (\frac{0.3}{1.45})^2 = 0.206896552^2 = 0.042806183$$

  So the probability is **4.28%**.

  Now, let's calculate it for $f^2(e)$:

$$P'(e_3 \cap e_3) = (P'(e_3))^2 = (\frac{f^2(e_3)}{\sum_{i=1}^{n} f^2(e_i)})^2 = (\frac{0.09}{0.7425})^2 = 0.121212121^2 = 0.014692378$$

  So the probability is **1.47%**.

3. What can be a good measure of performance for a genetic algorithm? Justify your answer! Using your measure find the optimal probability of the mutation for the backpacking problem, using elitism, a tournament selection with k = 4, a fitness function described in the first lecture (sum of the values if the sum of the weights is below or equal to the capacity, 0 otherwise). Is there a significant difference in the efficiency between the optimal parameter and setting the probability of mutation to 0?

   A good measure of performance for a genetic algorithm can be the approximation that it makes to the real optimum solution. The lower the difference is, the best performance the algorithm gives.

Bibliography

[1] Course Webpage
`http://math.bme.hu/ safaro/evolalgen.html`

[2] `https://tex.stackexchange.com/`

[3] `https://stattrek.com/probability/probability-rules.aspx`