technische universität
dortmund

Dept. of Computer Science
Chair of Algorithm Engineering

Tutorial for

Introduction to Computational Intelligence in Winter 2015/16

Günter Rudolph, Vanessa Volz

Lecture website: `https://tinyurl.com/CI-WS2015-16`

Sheet 6, Block III

28 January 2016

Due date: 10 February 2016, 2pm
Discussion: 11/12 February 2016

Exercise 6.1: Basic Probability Theory (4 Points)
Consider standard-bit-mutation on a bitstring of length n where the probability of flipping is $p = 1/n$ for each bit.

a) Calculate the probability that a certain bit is flipped at least once within t mutations.

b) Calculate the probability that exactly k bits of the bitstring are flipped in one mutation.

c) Given a bitstring x, calculate the probability that a certain bitstring y is the result of one mutation of x. Hint: Use the Hamming distance to relate bitstrings to each other.

d) Calculate the expected number of flipped bits per mutation.

Solution:

a) $1 - (1 - \frac{1}{n})^t$

b) $\binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{n-k}$

c) $\left(\frac{1}{n}\right)^{H(x,y)} \left(1 - \frac{1}{n}\right)^{n-H(x,y)}$

d) $\sum_{i=1}^{n} p_i = n\frac{1}{n} = 1$

Exercise 6.2: Real-valued Optimization (3 Points)
Download and install the R-package `cmaes` (`install.packages("cmaes")`). This package also contains the three test problems `f_sphere`, `f_rastrigin`, and `f_rosenbrock` (to be minimized). Compare the CMA-ES algorithm with the optimizer `rbga` in the package `genalg` (`install.packages("genalg")`) on these three problems. Repeat each algorithm at least 10 times for $n = 10$ decision variables. The number of function evaluations shall be fixed to 10000 and the search space be restricted to $[-5, 5]^n$. Average the obtained function values of the best solutions. Plot, report, and interpret the results.

Exercise 6.3: SMS-EMOA (9 Points)
You are playing a paladin in a video game and have three different attacks available. The mana needed to cast a spell and the damage caused are detailed in the table:

name	damage [HP]	mana [MP]
Avenger's Shield	4	1
Hammer of Wrath	3	1
Holy Nova	4	3

You have 5 mana available, but you can still only cast one spell once at most. Your goal is to select a combination of attacks that maximises the damage output while at the same time minimising mana consumption.

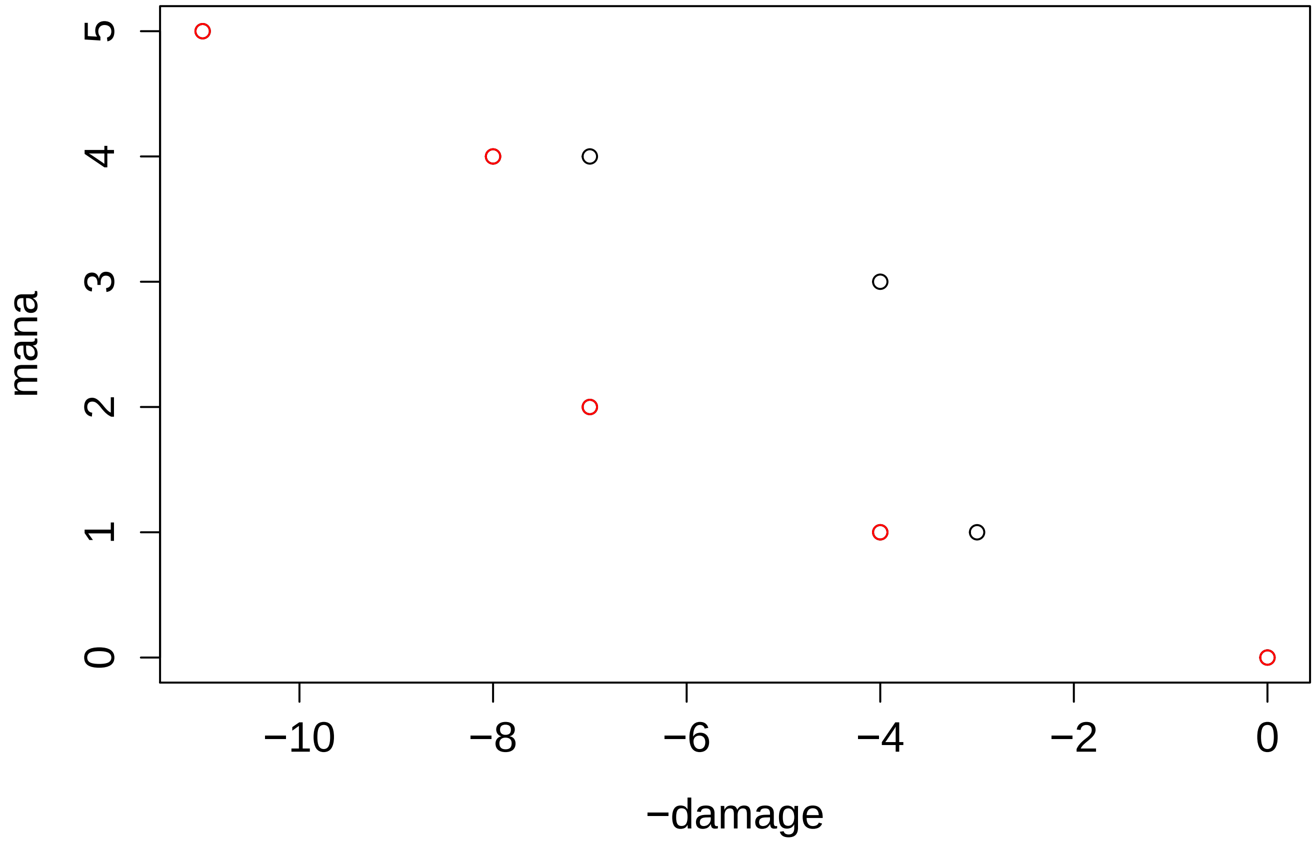

Figure 1: Solutions plot, non-dominated points are red

a) Plot the objective values of all possible solutions. Identify the non-dominated set. In a scenario with an opponent with 4 health points, which attack combination would you choose?

b) Assume you wanted to solve the problem as a single-objective minimisation problem and use a sum of the two objectives as a fitness value. Calculate the fitness value for all solutions. Discuss if there are any problems with aggregating the objective values with a weighted sum for this problem.

c) You now want to solve the problem as a multi-objective minimisation problem using the SMS-EMOA algorithm. Use a populations size of $\mu = 3$ with random initialisation, uniform crossover and global mutation with $p_m = \frac{1}{n}$ (n is the length of your genes). For a budget of 7 function evaluations calculate the result of the SMS-EMOA-run by hand. Detail all calculation steps and assume your random number generator results in the sequence of numbers in file randNo.txt.

Solution:

a) non-dominated set: (1,1,1) (1,0,1) (1,1,0) (1,0,0) (0,0,0), Scenario 4 HP: (1,0,0), plot in figure 1

b) The problem is (1,0,1) because with a weighted sum resulting in a linear function specifying the optimum, its value would never be found because together with the neighbouring non-dominated points, it forms a concave shape.

c)
- random initialisation, 1 if random value greater 0.5
  - (0,0,1) (rv 1-3)
  - (1,0,0) (rv 4-6)

    * (0,1,0) (rv 7-9)
  - function evaluations (3): (-4,3) (-4,1) (-3,1)
  - random parent selection: (1,3) (rv 10-11)
  - crossover (0,1,1) (rv 12-14)
  - mutation (0,1,1) (rv 15-17)
  - function evaluation (4): (-7,4)
  - Ranking: (1,0,0) and (0,1,1) have rank 1, others rank 2
  - Hypervolume contribution, refPoint is (1,6): hv(0,0,1)=3, hv(0,1,0)=8
  - resulting population: (1,0,0) (0,1,0) (0,1,1)
  - random parent selection: (1,2) (rv 18-19)
  - crossover (0,0,0) (rv 20-22)
  - mutation (0,0,1) (rv 23-25)
  - function evaluation (5): (-4,3)
  - child was eliminated before, so resulting population: (1,0,0) (0,1,0) (0,1,1)
  - random parent selection: (2,1) (rv 26-27)
  - crossover (1,1,0) (rv 28-30)
  - mutation (1,1,1) (rv 31-33)
  - function evaluation (6): (-11,5)
  - Ranking: only (0,1,0) has rank 2, rest 1
  - Resulting population (1,0,0) (0,1,1) (1,1,1)
  - random parent selection (1,3) (rv 34-35)
  - crossover (1,1,0) (rv 36-38)
  - mutation (0,1,1) (rv 39-31)
  - function evaluation (7): (-7,4)
  - child same as parent, so resulting population: (1,0,0) (0,1,1) (1,1,1)
  - stop after 7 function evaluations

Exercise 6.4: Variation operator Design (4 Points)

a) Approximate the entropy of the following distributions using their R implementations and verify that the normal distribution has the maximum value:

  - normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 4$: $N(0, 2)$
  - student-t distribution with $\frac{8}{3}$ degrees of freedom
  - laplace distribution with location $\mu = 0$ and scale $b = \sqrt{2}$

```r
dLaPlace = function(x, mu, b){
  return((1/(2*b))*exp(-abs(x-mu)/b))
}
```

Listing 1: `laPlaceDensity.R`

b) The parameters of the above functions are selected so that all distributions have a mean of 0 and a variance of 4. Why is that necessary when comparing the values of their entropy?

Solution:

- normal: 2.112086 with absolute error < 0.00014, analytic: $\frac{1}{2}\log(2\pi e\sigma^2) \approx 2.112086$
- student t: 1.819824 with absolute error $< 5.6 * 10^{-5}$, analytic: $\frac{df+1}{2}\left[\psi\left(\frac{1+df}{2}\right) - \psi\left(\frac{df}{2}\right)\right] + \log\left[\sqrt{df}B\left(\frac{df}{2}, \frac{1}{2}\right)\right] \approx 1.819824$, with digamma function ψ, beta function B and degrees of freedom df.
- la place: 2.039721 with absolute error $< 9.8 * 10^{-5}$, analytic: $\log(2be) \approx 2.039721$
- Normal distribution is max entropy if σ, θ are known, so the comparison in a) to verify that it is the maximum entropy distribution only makes sense for given σ, θ.