INF3490 exercises - week 1 2014

$\mathbb{P}$ marks the programming exercises.

Problem 1

Given the function $f(x) = -x^4 + 2x^3 + 2x^2 - x$:

$\mathbb{P}$ **a)** Plot the function, and its gradient, $f'(x) = -4x^3 + 6x^2 + 4x - 1$, in a plotting tool of your choice.

$\mathbb{P}$ **b)** Maximize using gradient ascent. You can try step size 0.1 and start somewhere in the range $[-2, 3]$. How does the choice of starting point and step size affect the algorithm's performance? Is there a starting point where the algorithm would not even be able to find a local maximum?

$\mathbb{P}$ **c)** Assume that we are only interested in maximums of $f(x)$ where x is between -2 and 3, and x increases in steps of length 0.5. Perform an exhaustive search to maximize $f(x)$ and plot the result.

d) In what way would greedy search and hill climbing differ for the maximization problem in Problem 2? Can you identify a starting position where the two algorithms might give different results?

e) Which algorithm do you think is the most efficient at maximizing $f(x)$ under the conditions in Problem 1c: exhaustive search or simulated annealing? Explain.

f) Gradient ascent, greedy search and hill climbing are quite similar, and are all based almost exclusively on exploitation. Can you think of any additions to these algorithms in order to do more exploration?

Problem 2

a) A common variant of evolution strategies used for (local) search is the $(1+4)$ ES. How would this differ from the $(1+1)$ ES in how the search space is explored? How does this, and $(1+\lambda)$ in general, compare to hill climbing and greedy search?

b) What effect does an adaptive search strategy have on optimization performance?

c) How would it affect the search if the strategy parameters were mutated after the solution parameters instead of before?

Problem 3

$\mathbb{P}$ **a)** Ignoring mutation, and starting with the population $\{1, 2, 3, 4\}$, implement and run 3 generations of a $(4+8)$ ES maximizing $g(x) = x$, and observe what the end population looks like (use intermediary recombination).

b) If an $(4, 8)$ ES had been used in Problem 3a, what would the probability of the optimal solution $(x = 4)$ surviving the first generation have been?

$\mathbb{P}$ **c)** Repeat Problem 3a with an EP with $q = 2$. How do the two algorithms compare?