# UNIVERSITY OF OSLO

## Faculty of Mathematics and Natural Sciences

**Exam in** **INF3490/4490 — Biologically Inspired Computing**
**Day of exam:** **December 3rd, 2014**
**Exam hours:** **14:30 – 18:30**
**This examination paper consists of 7 pages.**
**Appendices:** **1**
**Permitted materials:** **None**

*Make sure that your copy of this examination paper is complete before answering*.

**The exam text consists of problems 1-30 (multiple choice questions) to be answered on the form that is enclosed in the appendix and problems 31-33 which are answered on the usual sheets. Problems 1-30 have a total weight of 60%, while problems 31-33 have a weight of 40%.**

**About problem 1-30:**
Each problem consists of a topic in the left column and a number of statements each indicated by a capital letter. Problems are answered by marking true statements with a clear cross (X) in the corresponding row and column in the attached form, and leaving false statements unmarked. Each problem has a variable number of true statements, but there is always *at least one* true and false statement for each problem. 0.5 points are given for each marked true statement and for each false statement left unmarked, resulting in a score ranging from 0 to 60.

You can use the right column of the text as a draft. The form in the appendix is the one to be handed in (remember to include your candidate number).

**Problem 1**

| | | | |
|---|---|---|---|
| Hill climbing | A | Is a population-based optimization algorithm | |
| | B | Results depend on the starting points | |
| | C | Can only be done when a solution has a finite number of neighbors | |
| | D | Has less randomness than greedy search | |

**Problem 2**

| | | | |
|---|---|---|---|
| Strategy parameters | A | Adapt mutation using a fixed strategy schedule | |
| | B | Improve the chances of finding a better solution in the short term | |
| | C | Improve the chances of finding the global optimum | |
| | D | Adapt mutation by adjusting the normal distribution spread | |

**Problem 3**

| | | | |
|---|---|---|---|
| Evolution strategies have | A | Random parents selection | |
| | B | Uniform mutation | |
| | C | Recombination by partially mapped crossover | |
| | D | Fitness proportional survivor selection | |

**Problem 4**

| | | | |
|---|---|---|---|
| The crossover operators used in binary representations can also be used in | A | Integer representations | |
| | B | Real-valued representations | |
| | C | Permutation representations | |
| | D | Tree representations | |

**Problem 5**

| | | | |
|---|---|---|---|
| Permutation representation works with | A | Swap mutation | |
| | B | Creep mutation | |
| | C | Scramble mutation | |
| | D | Insert mutation | |

**Problem 6**

| | | | |
|---|---|---|---|
| Adding an offset to all fitness values affects selection pressure in | A | Fitness proportional selection | |
| | B | Ranking selection | |
| | C | Tournament selection | |
| | D | $(\mu + \lambda)$ selection | |

**Problem 7**

| | | | |
|---|---|---|---|
| One can improve results on multi-modal problems by | A | Ensuring that the initial population well distributed | |
| | B | Reducing the population size | |
| | C | Reducing the fitness of individuals that are close to others | |
| | D | Increasing the selection pressure | |

**Problem 8**

| | | | |
|---|---|---|---|
| Pareto dominance | A | Is hard to combine with tournament selection | |
| | B | Can be used to sort points according to multiple objectives | |
| | C | Reduces the objective functions to a scalar value | |
| | D | A solution dominates another if it is as good in every way and better in at least one | |

**Problem 9**

| | | | |
|---|---|---|---|
| Running multiple times is necessary to measure the performance of | A | An exhaustive search | |
| | B | An evolution strategy | |
| | C | Training a multi-layer perceptron | |
| | D | Training a self-organizing map | |

**Problem 10**

| | | | |
|---|---|---|---|
| Machine learning | A | Should be distinguished from self-learning | |
| | B | Is applicable to classification problems | |
| | C | A number of different biology-inspired methods could be used for machine learning | |
| | D | Is learning automatically from examples | |

**Problem 11**

| | | | |
|---|---|---|---|
| Machine learning | A | Can be applied to analyze new data | |
| | B | Is an alternative to artificial intelligence | |
| | C | Can be used at design time and/or at run time | |
| | D | Is always learning from scratch and not adaptation of a previously learned system | |

**Problem 12**

| | | | |
|---|---|---|---|
| Machine learning algorithms | A | Supervised learning is good for clustering problems | |
| | B | Reinforcement learning is about learning behavior based on reward | |
| | C | Unsupervised learning does not require target values | |
| | D | Selecting among the above learning methods is independent of the problem to be solved | |

**Problem 13**

| | | | |
|---|---|---|---|
| Swarm intelligence algorithms | A | Are inspired by interaction in nature between living beings in motion | |
| | B | Are focused on centralized control | |
| | C | Simple local rules are often applicable | |
| | D | It is difficult to predict the global behavior of the system | |

**Problem 14**

| | | | |
|---|---|---|---|
| Particle Swarm Optimization (PSO) | A | Is a population based algorithm | |
| | B | Particles are selected for survival based on their fitness | |
| | C | Velocity and position of each solution are updated | |
| | D | Updates are also based on neighbor particles | |

**Problem 15**

| | | | |
|---|---|---|---|
| Cartesian Genetic Programming (CGP) | A | Has less restrictions than Genetic Programming | |
| | B | Can be used for evolving digital circuits | |
| | C | The level-back parameter indicates the number of previous columns a node can connect to | |
| | D | Crossover is always used | |

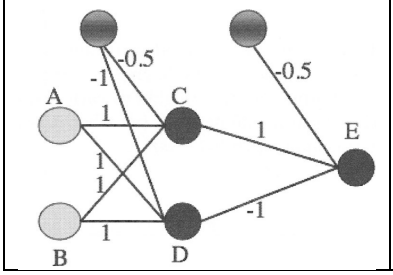
**Problem 16**

| | | | |
|---|---|---|---|
| Classification | A | Concerns finding decision boundaries that can be used to separate out different classes | |
| | B | Evolvable hardware is not applicable for classification | |
| | C | Non-linear decision boundaries can solve more complex problems than linear boundaries (straight lines) | |
| | D | A test set is more relevant for testing generalization than the training set | |

**Problem 17**

| | | | |
|---|---|---|---|
| Biological neural networks | A | The outputs from a neurons are pulses of fixed strength (height) and duration | |
| | B | The output from the neuron is called a synapse | |
| | C | Synapses can be inhibitory or excitatory | |
| | D | Learning takes place in the dendrites | |

**Problem 18**

| | | | |
|---|---|---|---|
| Which function does the following multi-layer perceptron realize:   | A | NAND | |
| | B | NOR | |
| | C | AND | |
| | D | XOR | |

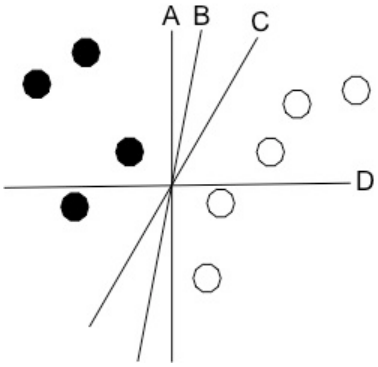
**Problem 19**

| | | | |
|---|---|---|---|
| Multilayer perceptron network | A | Usually, the weights are initially set to small random values | |
| | B | A hard limiting activation function is often used | |
| | C | The weights can only be updated after all the training vectors have been presented | |
| | D | Multiple layers of neurons allow for less complex decision boundaries than a single layer | |

**Problem 20**

| | | | |
|---|---|---|---|
| Support Vector Machines (SVMs) | A | Support vectors are used for computing hyperplanes | |
| | B | Is a method for minimizing the margin to hyperplanes | |
| | C | Nonlinear problems are handled with mapping inputs to lower-dimensional space | |
| | D | Kernel functions are used for transforming data | |

**Problem 21**

| | | | |
|---|---|---|---|
| Which separation line would SVM most likely choose?<br>  <br> | A | | |
| | B | | |
| | C | | |
| | D | | |

**Problem 22**

| | | | |
|---|---|---|---|
| Soft margins in SVMs | A | Reduce misclassifications during training | |
| | B | Allow some of the training data to be misclassified by introducing slack variables | |
| | C | Reduce the problem of training data overfitting | |
| | D | Are not useful if any training data is mislabeled | |

**Problem 23**

| | | | |
|---|---|---|---|
| Ensemble learning | A | A combination of classifiers are applied for classification | |
| | B | Classifiers should be trained to be slightly different | |
| | C | In bagging, each training sample (data point) is used only once for each iteration | |
| | D | Minority voting is used if there is disagreement | |

**Problem 24**

| | | | |
|---|---|---|---|
| Principal component analysis (PCA) | A | Finds the directions with the most variation in the data | |
| | B | Is useful for visualizing data | |
| | C | Dimensions are increased when applying PCA | |
| | D | Eigenvalues and eigenvectors are computed from the covariance matrix | |

**Problem 25**

| | | | |
|---|---|---|---|
| Unsupervised learning | A | Categorizes training vectors by identifying similarities between them | |
| | B | Can use the same error functions as supervised learning | |
| | C | Collaborative learning methods are often applied between classes | |
| | D | The data applied is unlabeled | |

**Problem 26**

| | | | |
|---|---|---|---|
| k-means | A | Automatically finds the number of clusters | |
| | B | Each cluster center is moved to the mean of data points assigned to it for each iteration | |
| | C | A too small number of clusters may lead to overfitting | |
| | D | The algorithm has converged when the change in cluster assignment is less than a threshold | |

**Problem 27**

| | | | |
|---|---|---|---|
| Self-Organizing Feature Map | A | Includes both a competition and collaboration part | |
| | B | Two or more weight layers are often used | |
| | C | Training data that are similar excite neurons that are near to each other | |
| | D | Represents a clustering technique | |

**Problem 28**

| | | | |
|---|---|---|---|
| Self-Organizing Feature Map learning | A | Increased network size leads to increased generalization | |
| | B | Weights of the winner neuron (and its neighborhood) are updated | |
| | C | The number of weights being modified for each training vector is increased throughout learning | |
| | D | A neighborhood function is used to compute the distance to the winner neuron | |

**Problem 29**

| | | | |
|---|---|---|---|
| Reinforcement learning | A | Works best with smaller state spaces | |
| | B | Keeps a log of all individual actions taken by the agent | |
| | C | Requires the agent to know the rewards for every action | |
| | D | Models learning behavior in animals | |

**Problem 30**

| | | | |
|---|---|---|---|
| Reinforcement learning discount factor | A | Is specified in the interval $[-1,0]$ | |
| | B | Is used to account for uncertainties about future rewards | |
| | C | Develops exponentially with time | |
| | D | Adjusts the balance between shortsightedness and farsightedness | |

**Problem 31 (6%)**

In a few sentences, sketch how you could modify a hill climbing algorithm in order to improve chances of finding the global optima.

**Problem 32 (10%)**

If you were to design an evolutionary algorithm to optimize the following problems, what kind of genetic representation (genotype) would you choose, and why? (Maximum two sentences for each)

a) Finding the best route for delivering a set of packages to different addresses
b) Optimize parameters of a physical structure like an antenna with a given shape
c) Design of a digital circuit

**Problem 33 (24%)**

SiO, the student welfare organization, would like to have a system for sorting utensils after washing. You are going to help them designing a camera based classifier system for sorting knifes, forks, spoons and teaspoons into separate bins. You have a machine vision library available that lets you identify where there is a utensil in the camera images, and it extracts a large number of features for each identified object that we can use as inputs.

(a) (4%) What class of learning algorithm would be best to use in this case, supervised, unsupervised or reinforcement learning? Justify your answer.
(b) (4%) We would like to make a system for distinguishing the utensils using a multi-layer perceptron network. How many output neurons should the network have, and what would each of them represent?
(c) (8%) Sketch the steps in the forward and backward phase of the multi-layer perceptron algorithm (backpropagation). Use words and not equations.
(d) (4%) What are the different approaches to how often weights are updated during training?
(e) (4%) How would you find out when to stop the training?

**INF3490/INF4490 Answers problems 1 – 30 for candidate no: _____**

| Problem | A | B | C | D |
|---|---|---|---|---|
| 1 | | | | |
| 2 | | | | |
| 3 | | | | |
| 4 | | | | |
| 5 | | | | |
| 6 | | | | |
| 7 | | | | |
| 8 | | | | |
| 9 | | | | |
| 10 | | | | |
| 11 | | | | |
| 12 | | | | |
| 13 | | | | |
| 14 | | | | |
| 15 | | | | |
| 16 | | | | |
| 17 | | | | |
| 18 | | | | |
| 19 | | | | |
| 20 | | | | |
| 21 | | | | |
| 22 | | | | |
| 23 | | | | |
| 24 | | | | |
| 25 | | | | |
| 26 | | | | |
| 27 | | | | |
| 28 | | | | |
| 29 | | | | |
| 30 | | | | |

**INF3490/INF4490 Answers problems 1 – 30**

| Problem | A | B | C | D |
|---|---|---|---|---|
| 1 | | O | | |
| 2 | | O | | O |
| 3 | O | | | |
| 4 | O | O | | |
| 5 | O | | O | O |
| 6 | O | | | |
| 7 | O | | O | |
| 8 | | O | | O |
| 9 | | O | O | O |
| 10 | | O | O | O |
| 11 | O | | O | |
| 12 | | O | O | |
| 13 | O | | O | O |
| 14 | O | | O | O |
| 15 | | O | O | |
| 16 | O | | O | O |
| 17 | O | | O | |
| 18 | | | | O |
| 19 | O | | | |
| 20 | O | | | O |
| 21 | | | O | |
| 22 | | O | O | |
| 23 | O | O | | |
| 24 | O | O | | O |
| 25 | O | | | O |
| 26 | | O | | O |
| 27 | O | | O | O |
| 28 | | O | | O |
| 29 | O | | | O |
| 30 | | O | O | O |