UNIVERSITY OF OSLO

Faculty of Mathematics and Natural Sciences

Exam in INF3490/4490 — Biologically Inspired Computing
Day of exam: November 29th, 2016
Exam hours: 09:00 – 13:00
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 1-35 (multiple choice questions) to be answered on the form that is enclosed in the appendix and problems 36-38 which are answered on the usual sheets (in English or Norwegian, please write clearly and sort sheets according to the problem numbers). Problems 1-35 have a total weight of 70%, while problems 36-38 have a weight of 30%.

About problem 1-35:
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. Further, -0.5 points are given for each marked statement not being true and for a correct statement not being marked. Thus, resulting in a score of max 70. If you think a statement could be either true or false, consider the most likely use/case.

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

Search	A	Exhaustive search is applicable for discrete problems	
	B	Greedy search makes the best choice available at each stage	
	C	Hill climbing compares the current best to all neighbours	
	D	Hill climbing is not applicable for continuous problems	

Problem 2

Which of the following are continuous optimization problems?	A	Prosthetic hand control	
	B	Timetable scheduling	
	C	Optimizing mechanical shapes	
	D	Prediction of stock prices	

Problem 3

Simulated annealing algorithm	A	Only improved solutions are kept during a run	
	B	The temperature is never increased	
	C	The search neighbourhood is increased during a run	
	D	Concerned with both exploration and exploitation	

Problem 4

Selection in evolutionary algorithms	A	Increases the diversity in the population	
	B	Implements competition between individuals	
	C	Pushes the population towards higher mean quality	
	D	Works on the individual level	

Problem 5

Recombination operators	A	Are not necessary if we use mutation	
	B	Usually include stochastic elements	
	C	Are used in every kind of evolutionary algorithm	
	D	Have to fit the genotypic representation	

Problem 6

Which variation operator(s) are applicable to permutation representations?	A	Swap mutation	
	B	Arithmetic crossover	
	C	Partially mapped crossover	
	D	1-point crossover	

Problem 7

Evolutionary algorithms (EAs)	A	Phenotypes and genotypes are usually identical	
	B	Selection operators need to be adapted to the genotypic representation	
	C	Fitness evaluation is applied to a phenotype	
	D	EAs are guaranteed to find the global optimum	

Problem 8

Selection operators	A	Fitness-proportionate selection may result in loss of selection pressure towards the end of runs	
	B	Rank-based selection is based on *relative* rather than absolute fitness	
	C	Tournament selection compares all individuals in the population	
	D	Uniform selection assigns the same probability of selecting every individual	

Problem 9

Survivor selection	A	(μ,λ)-selection is an elitist strategy	
	B	(μ,λ)-selection is better than $(\mu+\lambda)$-selection at leaving local optima	
	C	May be based on either age or fitness	
	D	$(\mu+\lambda)$-selection is an elitist strategy	

Problem 10

The simple genetic algorithm (SGA)	A	Does not use crossover	
	B	Does not use mutation	
	C	Can be used as a benchmark for new EAs	
	D	Uses a binary representation	

Problem 11

Problem variants	A	On-line control is a type of repetitive problem	
	B	Planning a daily mail delivery route is an example of a design problem	
	C	Evolutionary algorithms are not applicable for design problems	
	D	For design problems, we usually care most about *peak performance*	

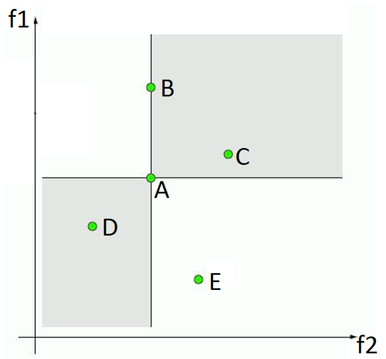
Problem 12

Multiobjective Evolution	A	Tries to approximate the Pareto front	
	B	Always relies on scalarization (taking a weighted sum) of the objectives	
	C	May use dominance relations to compare solutions	
	D	Only works if the objectives are not in conflict	

Problem 13

Which usually differ(s) between multiobjective and regular EAs?	A	The variation operator	
	B	The selection process	
	C	The diversification technique(s)	
	D	The genotypic representation	

Problem 14

Objectives f1 and f2 are both to be maximized. What is true about the plotted solutions? 	A	A dominates E	
	B	A dominates B	
	C	B dominates C	
	D	A and E do not dominate each other	

Problem 15

Supervised learning is appropriate for	A	Learning to play Atari games	
	B	Classification	
	C	Learning from unlabelled data	
	D	Regression	

Problem 16

Single-layer perceptrons	A	Can learn any function	
	B	Have exactly one hidden layer	
	C	Can be trained with supervised learning	
	D	Cannot be used for regression	

Problem 17

Multilayer perceptrons	A	Have one or more hidden layers	
	B	Only learn in the output layer	
	C	Are guaranteed to find the global optimum	
	D	Can be trained with the backpropagation algorithm	

Problem 18

Backpropagation	A	Requires pairs of input and target output	
	B	Uses the gradient descent technique	
	C	Does not require differentiable activation functions	
	D	Passes an error term *forward* through the network	

Problem 19

Neural network training	A	In batch training, weights are updated after each presentation of an input and target output	
	B	With minibatch training only one epoch is needed	
	C	A pass through all the training data is called an epoch	
	D	When training with a momentum, there is a higher chance of getting stuck in a local optimum	

Problem 20

Training and testing	A	Overfitting occurs when the model learns the bias in the training data	
	B	It is always best to train a classifier as long as possible	
	C	Test set is another name for validation set	
	D	We can avoid overfitting by increasing the size of a neural network	

Problem 21

Reinforcement learning	A	Requires pairs of input and target output	
	B	The goal is to maximize the total reward	
	C	Q-learning is an example of on-policy learning	
	D	When we have a Q-value, we do not require a policy	

Problem 22

SARSA	A	Iteratively updates its estimates of Q	
	B	Is an example of on-policy learning	
	C	Assumes we are following a greedy policy	
	D	Does not require a learning rate	

Problem 23

Deep Learning	A	May use the backpropagation algorithm	
	B	Requires features to be manually extracted from training data	
	C	Is neural networks learning with very many layers	
	D	Convolutional neural networks are appropriate for learning to classify images	

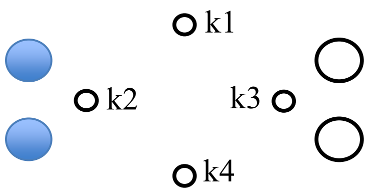
Problem 24

Unsupervised learning	A	Applicable for training with data sets containing only outputs	
	B	Reinforcement learning is applicable	
	C	Self–organizing maps are reducing the dimensionality of the data	
	D	Identifies clusters in the input data	

Problem 25

K–means clustering	A	A data point could belong to different clusters during a run	
	B	The number of clusters is being changed during a run	
	C	Each cluster center is moved least in the beginning	
	D	The method can result in a local minimum solution	

Problem 26

K–means clustering 	A	Cluster centers k1 and k2 would correctly distinguish the two classes (colored and white)	
	B	Cluster centers k1 and k4 would correctly distinguish the two classes (colored and white)	
	C	Cluster centers k2 and k3 would correctly distinguish the two classes (colored and white)	
	D	Using all four cluster centers would correctly distinguish the two classes (colored and white)	

Problem 27

Cartesian Genetic Programming	A	Mutation is less important than crossover	
	B	Is most commonly used for evolving computer programs	
	C	The level back parameter affects the extent of connections	
	D	The genome represents a non-regular and dynamic structure	

Problem 28

Particle Swarm Optimization	A	Works on a population of solutions	
	B	Generates new solutions by recombination of pairs of parents	
	C	Particles' updates depend on other particles in their neighbourhood	
	D	Uses the $(\mu+\lambda)$-selection strategy	

Problem 29

Support vector machines	A	Inputs are mapped into a lower-dimensional space	
	B	Is concerned with minimizing a margin	
	C	Kernel functions make separation of the data easier	
	D	Soft margins reduce the risk of overfitting	

Problem 30

Bagging	A	A method applicable to ensemble learning	
	B	Stands for bootstrap aggregation	
	C	A sample is taken from the original dataset with replacement	
	D	Each training vector is used once	

Problem 31

Boosting	A	Multiple classifiers are trained to be slightly different	
	B	Only the best classifier is applied after training	
	C	Training vectors are assigned weights during training	
	D	Misclassified training vectors are given lower weighs	

Problem 32

Dimensionality reduction	A	Increases the complexity of the training data	
	B	Principle component analysis is applicable	
	C	It could involve removing axes in the training data with least variation	
	D	Rotation matrices could be needed	

Problem 33

Uncanny valley	A	A challenging place in an optimization search space	
	B	An expression for when robots are very human-like	
	C	This may lead to people feeling a robot being a monster	
	D	A place with many robots being out of human control	

Problem 34

Reducing the risk of autonomous system misbehaving	A	Leave the human as much as possible out the loop at both design and run-time	
	B	Undertake thorough testing of the behaviour before applying it	
	C	Make the degree of autonomy dependent on the setting	
	D	Limit undesired access to control the system	

Problem 35

Recommendations for robots	A	Traceability depends on recording and documenting the robot behavior	
	B	Controlling and limiting a robot's autonomy can improve the identifiability	
	C	Password protection is important for privacy	
	D	Password protection is important for security and safety	

Problem 36 (9%)

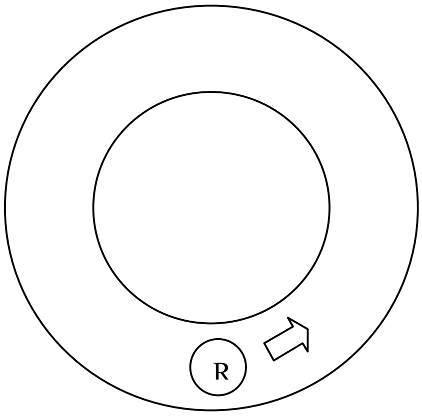
a) Briefly explain the terms *exploitation* and *exploration* related to search and how they differ.

b) Are search methods like greedy search and hill climbing most focused on exploitation or exploration? Explain why.

c) What additions to the methods in b) can be made to make them also cover the capability not covered so well (exploitation or exploration)?

Problem 37 (13%)

We would like to set up a neural network (multilayer perceptron) for robot control. The inputs are measurements from range sensors, and the output is a direction of movement. The robot is inserted into the circular maze shown to the right, and the goal is to enable it to drive in the direction of the arrow, getting as far as possible within a given time limit, while colliding with the walls as few times as possible.



a) One way to design this neural network is by use of an evolutionary algorithm (EA). The individuals in the population will be possible robot controllers that get their fitness computed in simulation. Assuming that the structure of the network is already specified, briefly describe how you could allow an EA to find the proper weights for this neural network. Include in your description a possible choice for:

a1) the genetic representation (genotype)

a2) variation operators. Include both their names and a *brief* description of how they work

a3) which measurements to include in the fitness function. You can assume the robot or the simulator can gather any physical measurements of relevance to fitness calculation.

b) A different way to solve this problem is to apply reinforcement learning (RL). Describe how you would model this problem as a reinforcement learning problem, including how you would define rewards, states and actions. The RL *algorithm* is *not* to be described.

Problem 38 (8 %)

a) Suppose the following set of points in two classes are to be distinguished using a Support Vector Machine:

class 1: (1,1), (2,0) and (3,1)

class 2: (1,4), (2,3) and (3,4)

Plot them and find the optimal separation line. Indicate what the support vectors are in the figure. What is the margin?

b) Would soft margins be beneficial for this data set? Justify your answer.

Appendix 1

INF3490/INF4490 Answers problems 1 – 35 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				
31				
32				
33				
34				
35				

INF3490/INF4490 Answers problems 1 – 35 for candidate no: __________

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

Problem 36 (9%)

a) Briefly explain the terms *exploitation* and *exploration* related to search and how they differ.

Exploration *is constantly* ***trying out completely new solutions*** *(global search).*
Exploration *is trying to* ***improve the current best solution*** *(local search).*

b) Are search methods like greedy search and hill climbing most focused on exploitation or exploration? Explain why.

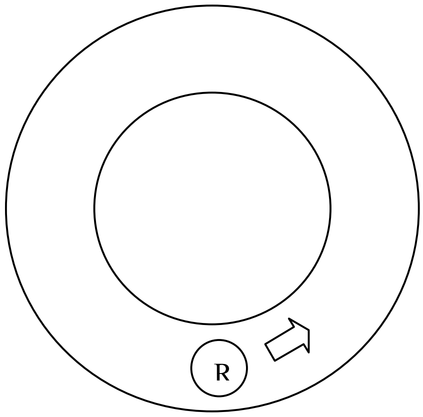
The methods are mostly focused on improving the current best solution, thus, ***exploitation****.*

c) What additions to the methods in b) can be made to make them also cover the capability not covered so well (exploitation or exploration)?

Run the algorithm several times *with random starting positions, this will explore the solution space and find several local optima. Another option is to* ***add more random movement*** *to either algorithm. This can be done after a solution is found, or at a probability while searching. Could also do backtrack + random jump after a solution is found.*

Problem 37 (13%)

We would like to set up a neural network (multilayer perceptron) for robot control. The inputs are measurements from range sensors, and the output is a direction of movement. The robot is inserted into the circular maze shown to the right, and the goal is to enable it to drive in the direction of the arrow, getting as far as possible within a given time limit, while colliding with the walls as few times as possible.



a) One way to design this neural network is by use of an evolutionary algorithm (EA). The individuals in the population will be possible robot controllers that get their fitness computed in simulation. Assuming that the structure of the network is already specified, briefly describe how you could allow an EA to find the proper weights for this neural network. Include in your description a possible choice for:

a1) the genetic representation (genotype)

a2) variation operators. Include both their names and a *brief* description of how they work

a3) which measurements to include in the fitness function. You can assume the robot or the simulator can gather any physical measurements of relevance to fitness calculation.

a1) *genetic representation (genotype): Since we are representing the weights of a neural network, the genotype needs to encode several numbers, that can be mapped to the neural network connections. The most straightforward way is to define each genotype as a list of floating-point values, where each value represents the weight of a single specific network connection.*

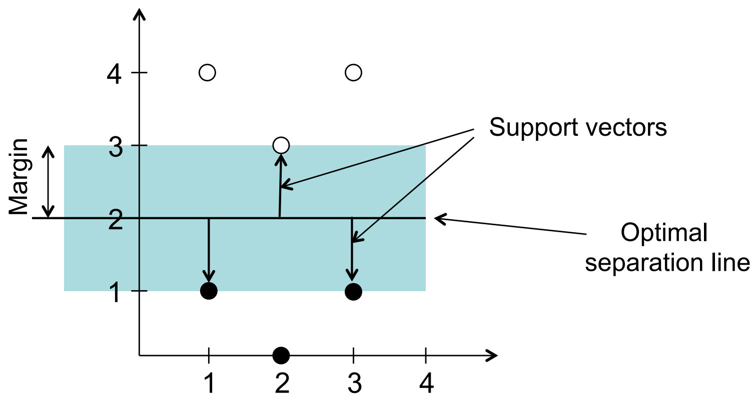
a2) *variation operators: Here, one should choose variation operators suitable for the representation defined in a1. Since we defined the genome as a list of floating-point values, we could for instance select uniform mutation and simple arithmetic crossover here. Other operators applicable to the representation are also accepted.*
a3) *fitness function: Since the goal of evolved controllers is to drive as far as possible within a time limit without crashing into walls, we should include measurements of the distance travelled and the total number of wall collisions in the fitness function.*

b) A different way to solve this problem is to apply reinforcement learning (RL). Describe how you would model this problem as a reinforcement learning problem, including how you would define rewards, states and actions. The RL *algorithm* is *not* to be described.

Since this robot control problem is continuous, rather than discrete, there is a potentially infinite number of different states and actions. We therefore need to ***discretize states and actions*** *before modelling this problem in the traditional RL way.*
For instance, we could model the problem this way:
States*: States need to include* ***information*** *about* ***distance to walls****. To guide the movements of the robot, we should also know on which side of the robot the wall is. There are many ways to represent this information. One example is to represent each state as two variables, one of which represents the* ***direction*** *towards the wall (dir), and the other the* ***distance*** *to it (dist). To guide actions, we need to discretize these states, for instance into the sets dir (left, front, behind, right) and dist (close, medium, far).*
Actions*: These need to be the* ***operations the robot can carry out*** *in order to complete its task. Again, we could discretize the robot's (continuous) control into a few different actions such as (go forward, go backward, turn left, turn right).*
Rewards*: These need to be adapted to the robot's goal, which is to drive far without collisions. For instance, one could give a* ***positive reward for every N cm driven****, and a* ***negative reward for every collision****.*

Problem 38 (8 %)

a) Suppose the following set of points in two classes are to be distinguished using a Support Vector Machine:
class 1: (1,1), (2,0) and (3,1)
class 2: (1,4), (2,3) and (3,4)
Plot them and find the optimal separation line. Indicate what the support vectors are in the figure. What is the margin?



b) Would soft margins be beneficial for this data set? Justify your answer.

No*, since the data set is easily separated with a linear line.*