6.034 Quiz 3
7 November 2014

Name	NASUS
Email	nasus@srift.net

Circle your TA **(for 1 extra credit point)**, so that we can more easily enter your score in our records and return your quiz to you promptly.

Josh Blum	**Elisa Castañer**	**Pedro Cattori**
Malcom Gilbert	**Dulcan Hamlet**	**Dylan Holmes**
Jessica Noss	**Duncan Townsend**	**Siyao Xu**

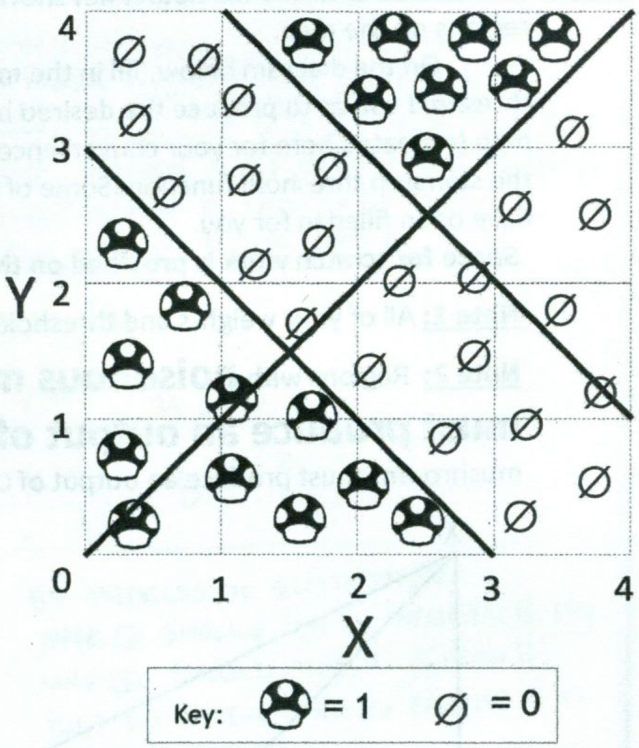
Problem number	Maximum	Score	Grader
1	50		
2	50		
Total	100		

SRN	6		

There are 11 pages in this quiz, including this one and including blank pages. As always, open book, open notes, open just about everything, including a calculator, but no computers.

Problem 1: Captain Teemo's Mushrooms (50 points)

The ground in Summoner's Rift includes dangerous areas full of poisonous mushrooms ( ), and safe areas without mushrooms (∅). In order to avoid stepping on poisonous mushrooms, Nasus constructs a map of these different areas (as shown on the right).

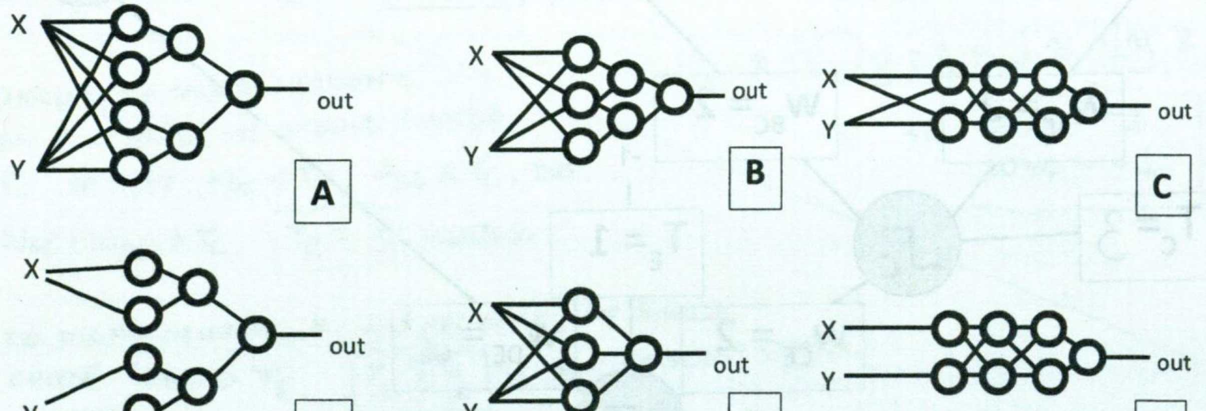


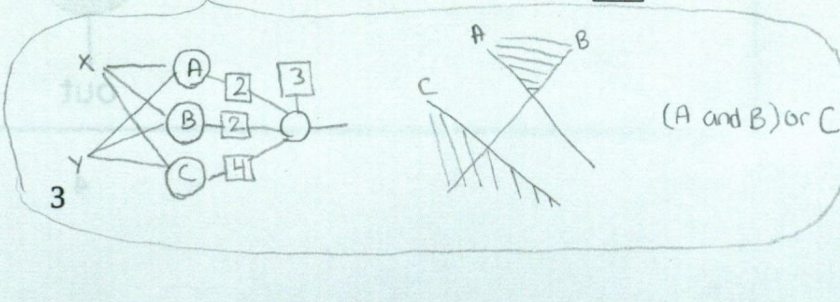
Nasus intends to use a neural net to predict whether a region has poisonous mushrooms or not.

Part A (18 points) Which of the following neural nets, if any, could be used to produce the classification regions shown in this map? (All neurons use the unit step threshold function; thresholds are not shown. Assume you can choose any values for the weights and thresholds.)

List **all** of the following neural nets (A, B, C, D, E, F) that can describe the regions in this map. If none of them can describe the regions in this map, write NONE OF THESE instead.

A, B, E

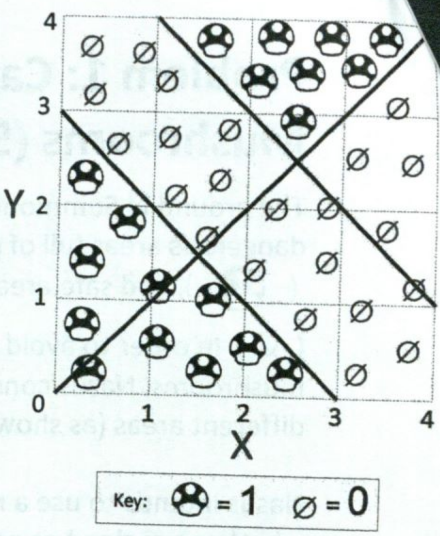




Part B (24 points) After considering several possible neural nets, Nasus decides to use the neural net shown below to produce the regions on the map.

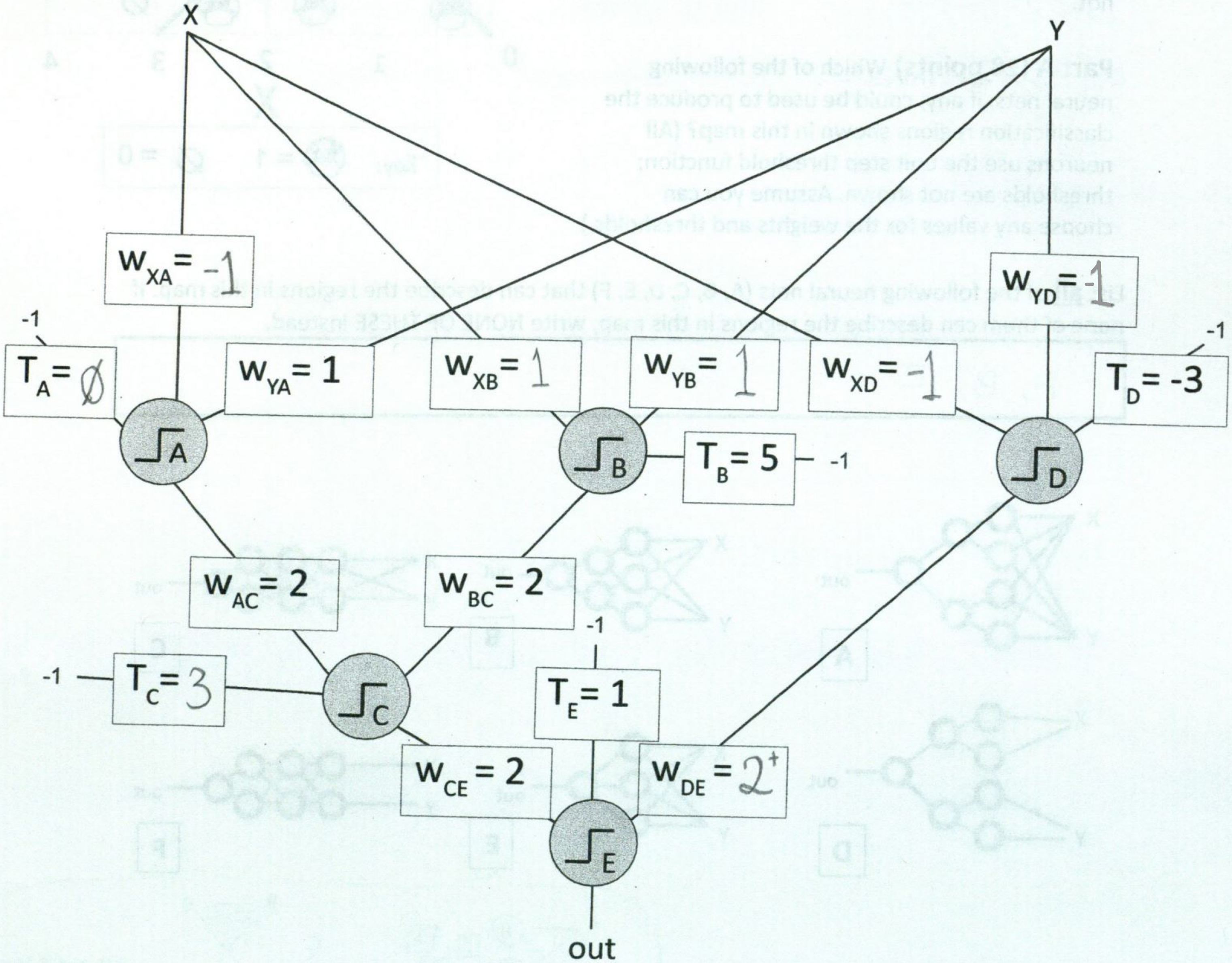
On the diagram below, fill in the missing weights and threshold values to produce the desired behavior as shown on the map (repeated here for your convenience). All of the neurons use the stairstep threshold function. Some of the weights/thresholds have been filled in for you.

Space for scratch work is provided on the next page.



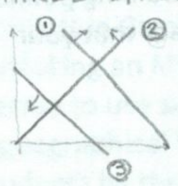
Note 1: All of your weights and thresholds must be **integers.**

Note 2: Regions with **poisonous mushrooms** **must produce an output of 1**; regions without mushrooms must produce an output of 0.



Space provided for scratch work.

THREE LINES



THE MOST OBVIOUS LOGIC FUNCTION IS TO HAVE (① AND ②) OR ③, SO:

- MAKE NEURON C INTO AN "AND" GATE
- ALIGN NEURONS A and B WITH LINES ① AND ②
- MAKE NEURON E INTO AN OR GATE
- ALIGN NEURON D AND LINE ③

LINES

① $y \geq 5 - x \Rightarrow \boxed{x + y \geq 5}$

② $y \geq x \Rightarrow \boxed{-x + y \geq \emptyset}$

③ $y \leq 3 - x \Rightarrow \boxed{-x + -y \geq -3}$

LINE-DRAWING NEURONS

$\boxed{A}$ ◯ $x + 1y \geq$ ◯

$\boxed{B}$ ◯ $x +$ ◯ $y \geq 5$

$\boxed{D}$ ◯ $x +$ ◯ $y \geq -3$

BECAUSE THE THRESHOLD OF LINE ② IS ZERO, IT ONLY MATCHES NEURON A; THE SCALING FACTOR IS +1 (NO SCALING REQUIRED)

$A = (+1) \times$ LINE ②

◯ $x + 1y \geq$ ◯ $= -1x + 1y \geq \emptyset$

SO WE HAVE $\boxed{W_{XA} = -1 \quad T_A = \emptyset}$

BY PROCESS OF ELIMINATION, LINE ① SHOULD MATCH NEURON B AND LINE ③ SHOULD MATCH NEURON D. THEY DO — THE SCALING FACTOR IS +1 IN BOTH CASES.

$B = (+1) \times$ LINE ①

◯ $x +$ ◯ $y \geq 5 = 1x + 1y \geq 5$

SO WE HAVE $\boxed{W_{XB} = 1 \quad W_{YB} = 1}$

$D = (+1) \times$ LINE ③

◯ $x +$ ◯ $y \geq -3 = -1x + -1y \geq -3$

SO $\boxed{W_{XD} = -1 \quad W_{YD} = -1}$

FINALLY, TO MAKE NEURON C AN AND GATE, WE SHOULD CHOOSE T_C SO THAT $W_{AC} < T_C$, $W_{BC} < T_C$, BUT $W_{AC} + W_{BC} > T_C$. $\boxed{T_C = 3}$ WORKS, as does $\boxed{T_C = 4}$.

TO MAKE NEURON E AN OR GATE, WE SHOULD CHOOSE $W_{DE} > T_E$. $\boxed{W_{DE} = 2}$ IS THE SIMPLEST CHOICE, but any $\boxed{W_{DE} \geq 1}$ is correct.

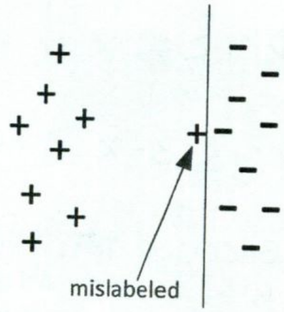
Part C (8 points)

① You train a neural net on a dataset consisting of six positive and six negative training points. Alas, after training, your neural net misclassifies half of the training points. The fact that your neural net misclassifies your training data in this way suggests that the issue is:

2

1. Overfitting
2. Underfitting
3. Cross-validation
4. The kernel trick

② While typing in data for an experiment with SVMs, you accidentally mislabel one of your training points. Consequently, the resulting SVM boundary has an extremely small margin. The susceptibility of SVMs to changes in just a few training points is known as:



1

1. Overfitting
2. Underfitting
3. Cross-validation
4. The kernel trick

③ Although both neural nets and support vector machines construct linear boundaries, only support vector machines construct the boundary with maximal margin width. What problem does maximization help to avoid?

1

1. Overfitting
2. Underfitting
3. Cross-validation
4. The kernel trick

④ For almost any training dataset, you can use the radial basis function to construct an SVM boundary that classifies all training points correctly—however, such a classifier may not accurately classify points outside of your training dataset. This problem is known as:

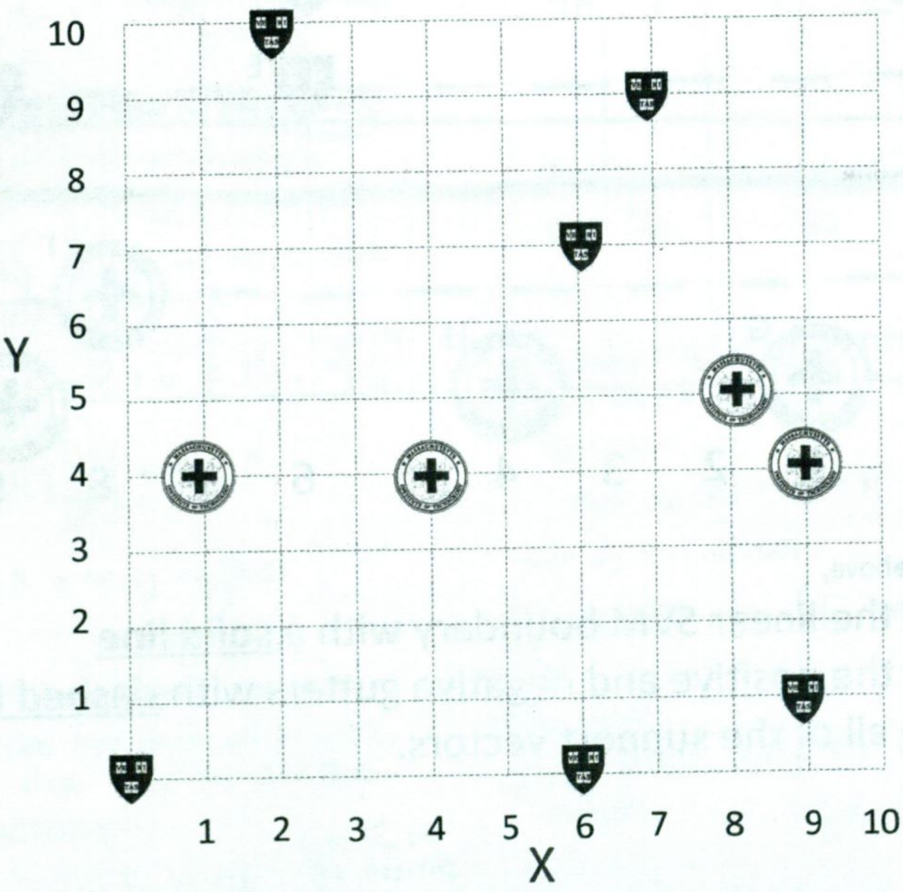
1

1. Overfitting
2. Underfitting
3. Cross-validation
4. The kernel trick

Problem 2: Support Vector Machines (50 points)

Tourists in Cambridge sometimes make the unforgivable blunder of mistaking an MIT student for a Harvard student and vice versa. You decide to use support vector machines to help them out—you'll paint a geographical boundary that separates MIT students and Harvard students by the widest possible margin. (Appropriately, you'll name your boundary *THE INFINITE SMOOT.*)

On the map below, you've charted the locations of several known MIT students (marked with a circle) and known Harvard students (marked with a shield).

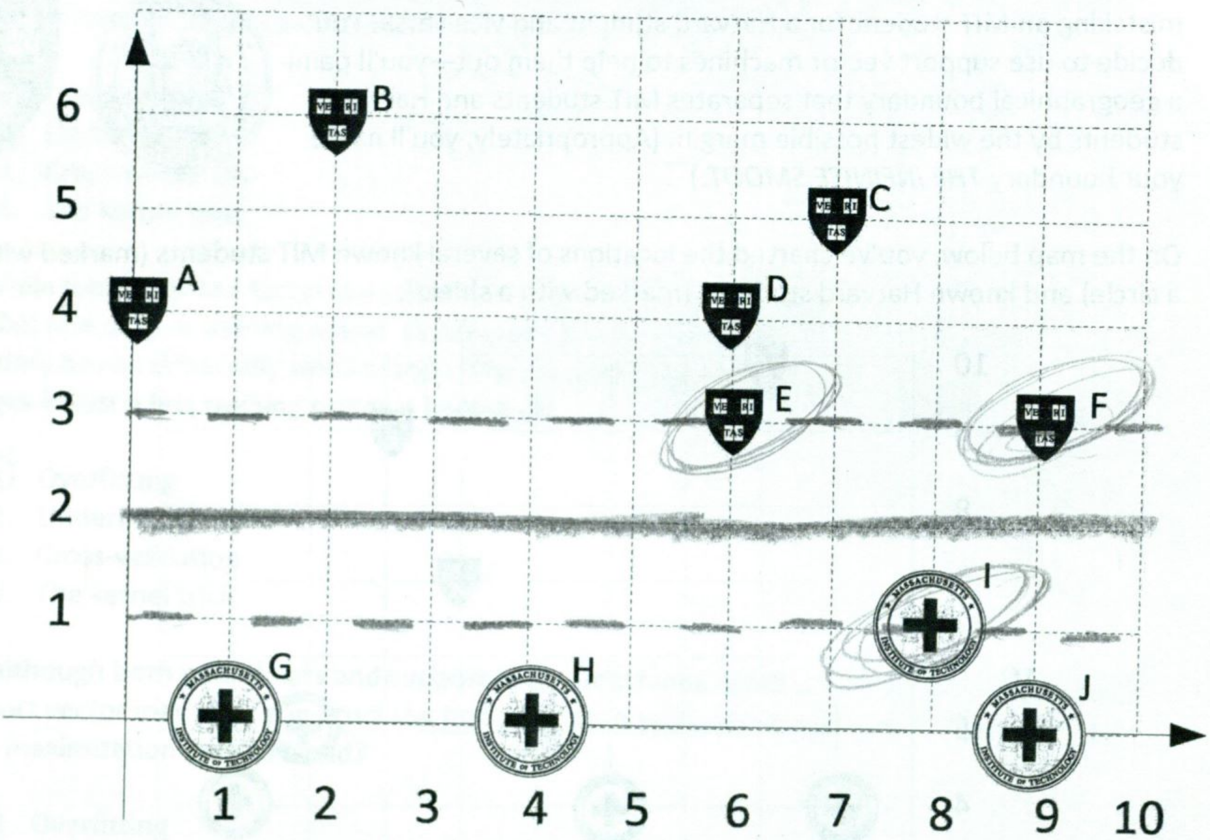


Part A (6 points) Which of the following kernel functions can be used to classify the students in the above dataset? (Circle **all** that apply. If none of them can classify this dataset, circle NONE OF THESE instead.)

LINEAR KERNEL   QUADRATIC KERNEL   RADIAL BASIS KERNEL

NONE OF THESE

Part B (12 points) To make classification easier, you decide to **transform your dataset** to make it linearly separable. Below is a map showing the transformed dataset. The points have now been labeled for your convenience.



On the diagram above,

- Draw the linear SVM boundary with a <u>**solid line**</u>
- Draw the positive and negative gutters with <u>**dashed lines**</u>.
- <u>**Circle**</u> all of the support vectors.

Part C (22 points) Based on the boundary and gutters you drew in Part B, determine the values of $\vec{w}$, b , and the α_i for the boundary equation $\vec{w}\cdot\vec{x}+b=0$ in this transformed space. Your classifier must produce **an output of +1 for MIT students**, and an output of -1 for Harvard students.

= +1	= -1

$\vec{w} = [\emptyset, -1]$ $\alpha_A = \emptyset$ $\alpha_B = \emptyset$ $\alpha_C = \emptyset$ $\alpha_D = \emptyset$ $\alpha_E = 1/6$

$b = 2$ $\alpha_F = 1/3$ $\alpha_G = \emptyset$ $\alpha_H = \emptyset$ $\alpha_I = 1/2$ $\alpha_J = \emptyset$

Show your work below for partial credit.

THE EQN OF THE BOUNDARY IS

$$Y \leq 2$$

WHICH WE RELATE TO THE GENERAL EQUATION

$$\vec{w}\cdot\vec{x} + b \geq \emptyset$$ ON THE POSITIVE SIDE.

$$[w_1 \;\; w_2]\begin{bmatrix} x \\ y \end{bmatrix} + b \geq \emptyset$$

To account for scaling, we multiply by a constant $c > \emptyset$

$$c[w_1 x + w_2 y + b] \geq \emptyset$$

$$-y + 2 \geq \emptyset$$

SO $w_1 = 0$ AND $w_2 = -1$ AND $b = 2$

TO FIND THE SCALING FACTOR, WE CAN USE THE GUTTER CONSTRAINT

$$\vec{w}\cdot\vec{x}_+ + b = +1$$ FOR $\vec{x}_+$ ON THE POSITIVE GUTTER

$$[cw_1 \;\; cw_2]\cdot\begin{bmatrix} 8 \\ 1 \end{bmatrix} + cb = 1$$

$$[\emptyset \;\; -c]\cdot\begin{bmatrix} 8 \\ 1 \end{bmatrix} + 2c = 1$$

$$\emptyset\cdot 8 + (-c)(1) + 2c = 1$$

$$\boxed{c = 1}$$

SO NO SCALING IS REQUIRED.

$$\vec{w} = [cw_1 \;\; cw_2] = [\emptyset \;\; -1]$$

$$b = 2c = 2$$

TO FIND THE α_i, WE FIRST NOTICE THAT $\alpha_i = \emptyset$ FOR NON-SUPPORT VECTORS:

$$\alpha_A = \alpha_B = \alpha_C = \alpha_D = \emptyset = \alpha_G = \alpha_H = \alpha_J$$

THEN WE CAN DEPLOY

$$\sum_{\text{SUPPORT}} y_i\alpha_i = \emptyset \quad \text{SO} \quad \alpha_I - \alpha_E - \alpha_F = \emptyset \qquad (1)$$

AND ALSO

$$\vec{w} = \sum_{\text{SUPPORT}} y_i\alpha_i\vec{x}_i \quad \text{SO} \quad 8\alpha_I - 6\alpha_E - 9\alpha_F = \emptyset \qquad (2)$$

$$1\alpha_I - 3\alpha_E - 3\alpha_F = -1 \qquad (3)$$

$$\tfrac{1}{2}[(3) - (1)] : \alpha_I = 1/2 \qquad (4)$$

$$(4) - (1) : \alpha_E + \alpha_F = 1/2 \qquad (5)$$

$$-8(1) + (2) : 2\alpha_E - \alpha_F = \emptyset \qquad (6)$$

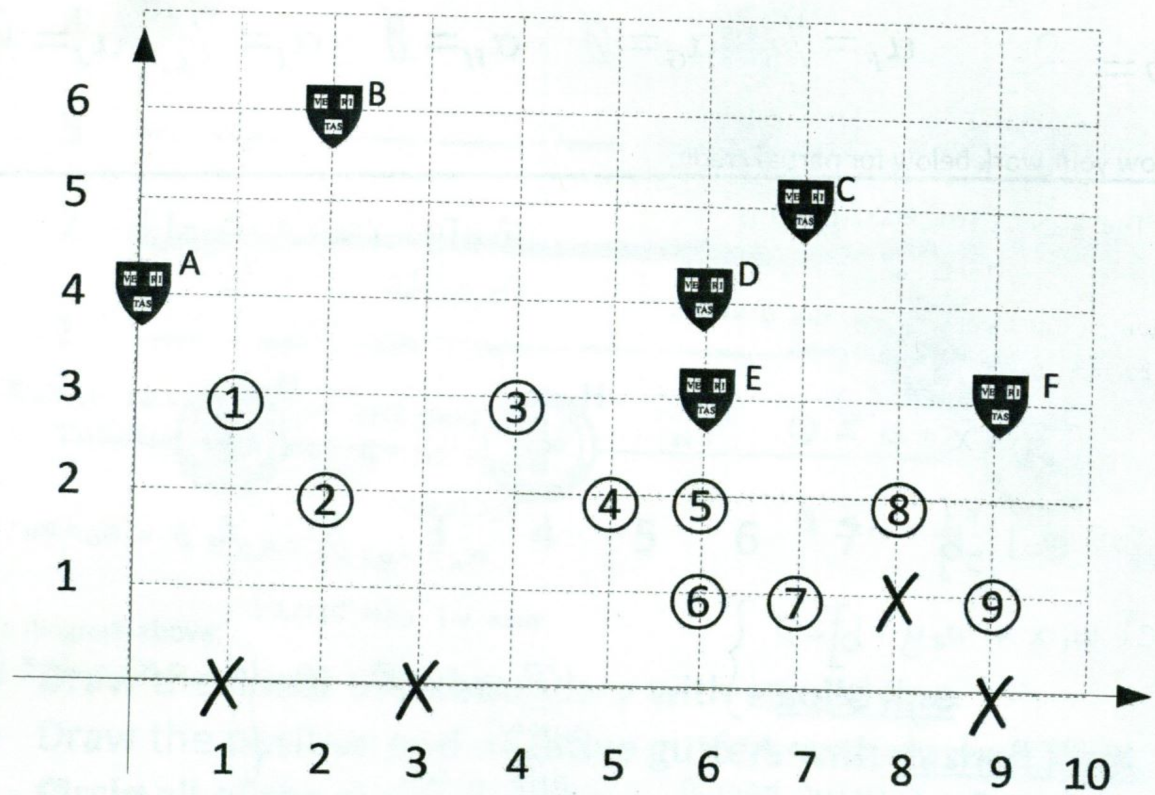
$$\tfrac{1}{3}[(5) + (6)] : \alpha_E = 1/6 \qquad (7)$$

$$\tfrac{1}{3}[2(5) - (6)] : \alpha_F = 1/3 \qquad (8)$$

Part D (10 points) Suppose all of the MIT students are removed from the training set and replaced with a single MIT example—say, Professor Winston. Where could Prof. Winston stand by himself so that **the boundary from Part B would not change**?

(For your convenience, the former location of each deleted MIT training point is marked with an X.)

<u>**Reminder**</u>: Since the former MIT training points have been deleted, they have no effect on the boundary anymore.



List **<u>all</u>** of the places ①②③④⑤⑥⑦⑧⑨ where Prof. Winston could stand so that the boundary will be the same as it was in Part B. If Prof. Winston can't stand in any of those places without changing the boundary, write NONE OF THESE instead.

6, 7, 9

Problem 3: Spiritual and Right-Now (6 points)

Circle the **best** answer for each of the following questions. There is **no penalty for wrong answers,** so it pays to guess in the absence of knowledge.

1 Detailed study of drawing behavior indicates:

1

1. The disappearance of hooklets indicates onset of mental impairment.
2. The appearance of hooklets indicates onset of mental impairment.
3. Victims of advanced Alzheimer's can draw emoticons but cannot name them.
4. Victims of advanced Alzheimer's can name emoticons but cannot draw them legibly.
5. Instrumented pens are helpful in Alzheimer's testing because they reveal minute vibrations.

2 Recent evidence from fMRI indicates

5

1. No specialized area for any visual categories unknown 100,000 years ago.
2. No specialized area for any visual category other than faces, places, and body parts.
3. Specialized areas for visual categories such as predators and prey.
4. A specialized area that responds to basic actions such as dropping and picking up.
5. A specialized area for words.

3 Genetic algorithms use crossover:

3

1. To make small adjustments to mutation-based hill-climbing rate.
2. To refine the mapping from genotype to phenotype.
3. To make large jumps in genotype space.
4. To bridge the gap between phenotype and fitness.
5. To add a diversity metric to fitness calculations so as to fill evolutionary niches.

4 Learning phonological rules using the Sussman-Yip method:

5

1. Uses the require-link, forbid-link, and jump-link heuristics.
2. Uses a kernel function to transform phoneme space into a sparse space.
3. Determines which phonological rule applies using Hamming distance on distinctive features.
4. Evolves a recognition pattern using boosting.
5. Forms its own negative examples by exploiting the fact that words have only one plural form.

5 Cross-training refers to:

2

1. Multiple humans training a robot so the robot has multiple ways of doing its tasks.
2. Robots and workers exchanging roles so as to develop better mental models of each other.
3. Preparation for robotics research by studying both cognitive science and computer science.
4. Teaching robots in a simulated world to do work in the real world.
5. Teaching robots by using human teachers already trained to do the work.

6 Architectures that use chunking:

5

1. Arrange processing modules in hierarchical groups.
2. Use wrappers to combine multiple values into propagated objects.
3. Find heuristically relevant methods using nearest neighbor calculations.
4. Divide problems into "chunks" handled by specialized processors.
5. Remember useful sequences of actions.