In the below table of data you can see the features of a number of a animals and a Boolean variable (label) that indicates if each animal in the dataset is a dog or not. Using a linear binary classification model, there will be a weight vector with two dimensions that can be used as, g(X)= W.X+b= w1x1+ w2x2+b and the decision function will be the Sign(g(x)). However, you can see this simple binary classification problem as a structured learning problem. In this case, you will have a weight vector for each class label here assume the Dog and NotDog classes. The learning function can be written in terms of both x and y and the decision will be made by finding the answer to the following:

$$h(x; W) = \arg\max_{y \in \mathcal{Y}} g(x, y; W)$$

A) find the g(x,y;W) for the first example, assume $W = \begin{bmatrix} 0.7 \\ 0.1 \\ -0.1 \\ 0.9 \end{bmatrix}$

(First two dimensions are for dog class and the other two are for notdog class. slide 15-StructuredMachineLearning slide#6 might be helpful, same thing in slide 14-StructuredMachineLearning, slide#8.)

	Barks	Color-brown	Dog
	1	1	1
	0	1	0
	1	0	1
	0	1	0

B) Assume that you still want to continue training your model to improve the weights. In the loop of training we are using the loss-augmented inference as follows:

$$\arg\max_{y \in \mathcal{Y}} (g(x^i, y; W) - g(x^i, y^i; W) + \Delta(y^i, y))$$

This equation helps you to find the most violated label for each training example. You want to choose the most violated labels [dog, notdog] for the input example (Barks=1, Color-brown=1), the first input example x with a number of possible output labels. Given the above equation, how many outputs are possible and which one is the most violated one?