UNIVERSITY OF OSLO

Candidate no

Faculty of mathematics and natural sciences

Exam in:	TEK5040 — Deep Learning for Autonomous Systems
Day of examination:	6th December 2019
Examination hours:	09:15 – 13:15

This problem set consists of 5 pages.

Appendices:	None.
Permitted aids:	None.

Please make sure that your copy of the problem set is complete before you attempt to answer anything.

(Continued on page 2.)

Contents

Problem 1 Semantic segmentation (weight 7%)

Assume that you are training a segmentation model to segment out the lines that mark the different lanes on a road. Everything except theses lines are treated as the background class. Assume that you obtain 99 % accuracy on the test set. What is an undesirable property of accuracy as a measure of performance in this example? Give at least two examples of alternative metrics that can be used.

Problem 2 Image captioning with attention (weight 12%)

Assume that you want to train a Recurrent Neural Network (RNN) to generate captions for images, and that you have a pretrained CNN that extracts feature maps of dimension $h \times w \times c$ from the images, where h, w and c are the height, width and number of channels of the feature maps. Describe one way you could use content-based attention to focus on a particular part of the image at each time step.

Problem 3 Greedy policy update (weight 7%)

Assume that we have a policy π and let v_π and q_π denote the state-value and action-value function respectively. Assume a discrete action space. Define a new policy π' by

$$\pi'(s) := \text{argmax}_a q_\pi(s, a)$$

i.e. greedily choose the action that looks the most promising. Is it possible that the new policy is *worse* for some states, i.e. that $v_{\pi'}(s) < v_\pi(s)$ for some state s? Justify your answer.

(Continued on page 3.)

Problem 4 Proximal Policy Optimization (PPO) (weight 10%)

Let π_θ denote the policy parametrized by $\theta \in \Theta$. Assume that we have sampled a number of episodes using the weights θ_{old} for our policy. Define

$$u_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)} \tag{1}$$

and let $\hat{d}_t$ denote the estimated advantage of taking action a_t from state s_t. The PPO surrogate objective is then defined as

$$L^{PPO}(\theta) = \hat{E}_t[\min\left(u_t(\theta)\hat{d}_t, \text{clip}(u_t(\theta), 1-\epsilon, 1+\epsilon)\hat{d}_t\right)] \tag{2}$$

where $\epsilon > 0$ is a hyperparameter and $\hat{E}_t$ denotes the expectation over the sampled dataset.

What is the purpose of the hyperparameter ϵ used in clip operation?

Problem 5 Meta learning (weight 12%)

Tom plans to train an image classifier with a training set which contains many images of 100 classes. He also has a test set consisting of images of all the 100 classes. Bob gets this classifier and tries to use it to classify a given image into one of the three new classes {*truck, train, motorcycle*}. Unfortunately, Bob has only one example of each class. Bob and Tom decide to use meta learning with the Matching network architecture to tackle the problem.

5a (weight 5%)

If the task of Tom was a regular training, a data sample would be $(\boldsymbol{I}, \boldsymbol{c})$, where $\boldsymbol{I}$ is a sample image and $\boldsymbol{c}$ is its class label. In meta learning, what is the structure of a training data sample *Tom* can use? (Hint: Think of a meta-training set)

5b (weight 4%)

What is the structure of a data sample *Bob* can use? (Hint: Think of a meta-test set)

5c (weight 3%)

The Matching network outputs a probability vector $[p_1, p_2, p_3]$. What do p_1, p_2 and p_3 measure?

(Continued on page 4.)

Problem 6 Processing 3D-Scenes (weight 12%)

6a (weight 2%)

One approach for segmentation and classification of 3D scenes is to apply 3D-convolutions on voxelized point clouds. Briefly describe a problem with this approach.

6b (weight 4%)

Give a short explanation of how the Pointnet achieves *permutation invariance* in processing point clouds.

6c (weight 6%)

Using graph convolutions, we can exploit the information of local structure in point clouds. A general graph convolution called *edge convolution* can be defined using the function $\boldsymbol{h}_\theta(\cdot,\cdot)$ which is parameterized by θ:

$$\boldsymbol{x}'_i = \mathop{\square}_{j\,:\,(i,j)\,\in\,\mathcal{E}} \boldsymbol{h}_\theta(\boldsymbol{x}_i, \boldsymbol{x}_j),$$

where $\square$ denotes the aggregation operation for the node i over all its neighbours j, whereas $\boldsymbol{x}$ and $\mathcal{E}$ denote feature vectors and the set of all edges in the graph respectively. Show that the regular convolution used in CNNs and operations used in the Pointnet are special cases of this edge convolution.

Problem 7 Tracking in video (weight 11%)

7a (weight 2%)

Name four basic operations in multiple object visual tracking (MOT) in which deep learning can be employed.

7b (weight 4%)

Why can a siamese network be a simpler and more efficient way of tracking objects in a video compared to the MDNet?

7c (weight 5%)

A problem with siamese networks is that they do not handle different scales of the tracked object well. Give an outline of a method with which we can improve on that.

(Continued on page 5.)

Problem 8 Conditional GAN (weight 10%)

Assume that you have a set of color images $x_1, x_2, \ldots, x_N$ and you want to train a neural network to *colorize* grayscale images, i.e. so that given a grayscale image x' the network may generate different plausible color images. As an application, you could then use this network to possibly colorize old digitized grayscale photos. Briefly describe how you can formulate this as a conditional GAN problem.

Problem 9 Evidence lower bound (weight 7%)

9a (weight 3%)

In variational inference we try to approximate the posterior distribution $p(\boldsymbol{w}|\mathcal{D})$ with an auxiliary distribution $q(\boldsymbol{w})$, where $\boldsymbol{w}$ and $\mathcal{D}$ are network parameters and training data respectively. The KL divergence between the two distributions is given by

$$\mathrm{KL}\big(q(\boldsymbol{w})||p(\boldsymbol{w}|\mathcal{D})\big) = -\mathbb{E}_{q(\boldsymbol{w})} \ln \frac{p\left(\boldsymbol{w}, \mathcal{D}\right)}{q\left(\boldsymbol{w}\right)} + \ln p(\mathcal{D})$$

Identify the quantity known as Evidence Lower Bound (ELBO) in the above equation.

9b (weight 4%)

Explain why maximizing ELBO would lead the distribution $q(\boldsymbol{w})$ to get closer to $p(\boldsymbol{w}|\mathcal{D})$.

Problem 10 Deep learning for control (weight 12%)

10a (weight 2%)

What is the main motivation for inverse reinforcement learning?

10b (weight 5%)

Assume that you are given a set of expert demonstrations together with a policy network and a reward network initialized to random values. The Guided Cost Learning (GCL) algorithm is used to train the policy and reward networks. Write down a very high level pseudo code of the GCL algorithm. (Hint: Show only the three basic steps of the algorithm).

10c (weight 5%)

Modify your pseudo code to that of Generative Adversarial Imitation Learning (GAIL).