Quiz 2
December 2, 2016

Time: 50 minutes; Total Marks: 45
One double-sided 8.5" x 11" cheat sheet allowed

This test contains 3 questions and 7 pages

NAME: _____

STUDENT NUMBER: _____

Question	Marks	Time budget
1	/24	25 min
2	/9	10 min
3	/12	10 min

1. (24 marks) True or False questions. **Provide a short explanation.**

   (a) True or False. If $k_1(\boldsymbol{x}, \boldsymbol{z})$ and $k_2(\boldsymbol{x}, \boldsymbol{z})$ are valid kernels, then so is $k_3(\boldsymbol{x}, \boldsymbol{z}) = k_1(\boldsymbol{x}, \boldsymbol{z}) + k_2(\boldsymbol{x}, \boldsymbol{z})$.

   (b) True or False. At test time, evaluating a kernelized perceptron on an input $\boldsymbol{x}$ is faster than evaluating an equivalent perceptron using an explicit basis function representation ($\boldsymbol{w}^T \boldsymbol{\phi}(\boldsymbol{x})$).

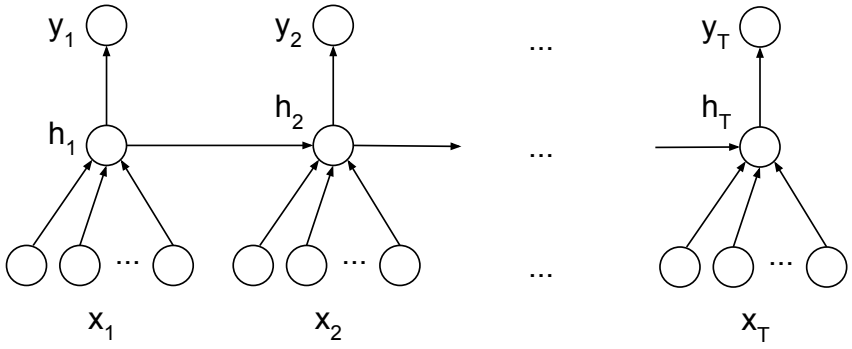
   (c) True or False. For random variables A, B, and C, $P(A, B|C) = P(A|C)P(B|C)$.

(d) True or False. Any potential function defined over 3 variables can be written as a product of potential functions over pairs of variables.
I.e. $\psi_{ABC}(A, B, C)$ can be written as $\psi_{AB}(A, B) \cdot \psi_{BC}(B, C) \cdot \psi_{AC}(A, C)$.

(e) True or False. If a first order Hidden Markov Model is used for object tracking, the object's acceleration is not modeled.

(f) True or False. If we use a recurrent neural network model for speech recognition, the output word at time t depends only on the sound heard at time t.

2. (9 marks) Consider the recurrent neural network shown below.



Use the notation below:

- Input (vector) at time t is $\boldsymbol{x}_t$
- Weights (vector) connecting input to hidden unit $\boldsymbol{w}_{hx}$
- Weight connecting previous hidden unit to current w_{hh}
- Activation function for hidden unit h_t is $\phi(\cdot)$
- Weight connecting hidden unit to output w_y
- Activation function for output unit y_t is $\sigma(\cdot)$

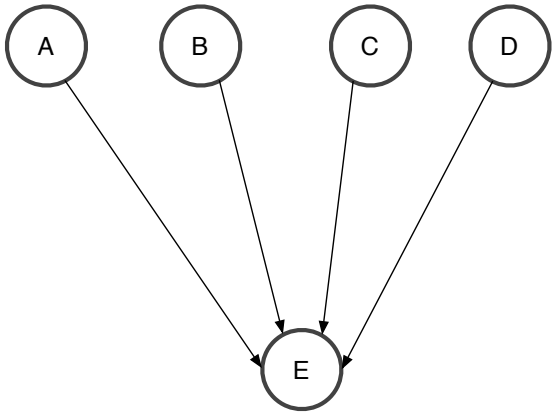
(a) Write down a mathematical formula for what the output y_1 will be, given an input sequence $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T)$

(b) Write down a mathematical formula for what the output y_2 will be, given an input sequence $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T)$

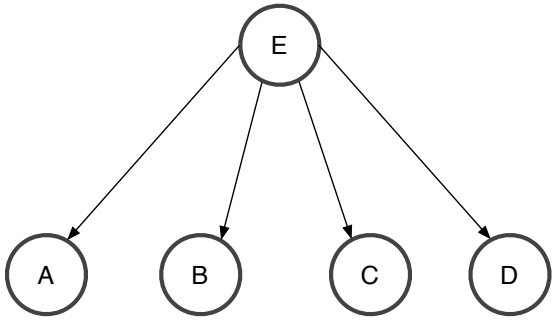
(c) Which inputs $\boldsymbol{x}_t$ does an output y_k depend on?

3. (12 marks) Answer the questions about the graphical models below. **Assume the random variables A, B, C, D, and E are binary.**

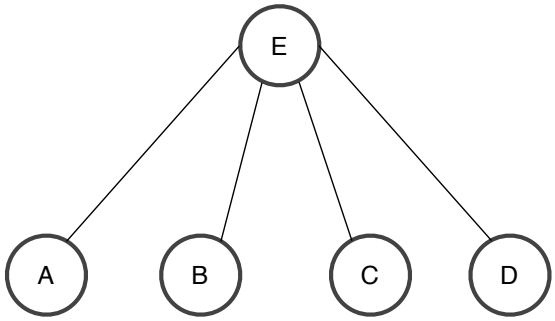
   (a) How many parameters are needed to specify the probability distributions in the Bayesian network below?



   (b) How many parameters are needed to specify the probability distributions in the Bayesian network below?



(c) How many parameters are needed to specify the potential functions in the Markov Random Field below?



(d) Do the Bayesian network and Markov Random Field below have the same conditional independence assumptions? Why or why not?