CS3236: Midterm Exam

20 October 2014

Name:
Student Number:

This is a **closed book** exam. The duration of the test is 1 hour and 30 minutes and consists of 30 questions. You may **not** use computers or electronic computers during the test.

For each question, one and only one answer is correct. Answer by circling the item you choose. If you answer correctly, you get 3 points. If you answer wrongly, you get 1 point penalty. This means you should not guess at random the answer to a question. The highest possible mark is 90 points. Given the time constraint, allocate around 3 minutes per question.

1. What is the definition of conditional entropy?

   a) $H(X \mid Y) := -\sum_{x,y} \Pr_X(x) \Pr_{X|Y}(x \mid y) \log \Pr_X(x)$
   b) $H(X \mid Y) := -\sum_{x,y} \Pr_X(x) \Pr_{X|Y}(x \mid y) \log \Pr_{X|Y}(x \mid y)$
   c) $H(X \mid Y) := -\sum_{x,y} \Pr_Y(y) \Pr_{X|Y}(x \mid y) \log \Pr_Y(x)$
   d) $H(X \mid Y) := -\sum_{x,y} \Pr_Y(y) \Pr_{X|Y}(x \mid y) \log \Pr_{X|Y}(x \mid y)$

2. Which of the following inequality is always true, whatever random variables X, Y, and value y?

   a) $H(X \mid Y) \leqslant H(X)$
   b) $H(X \mid Y = y) \leqslant H(X)$
   c) $H(X \mid Y) \geqslant H(X)$
   d) $H(X \mid Y = y) \geqslant H(X)$

3. What is the definition of mutual information?

   a) $I(X;Y) := H(Y) - H(X \mid Y)$
   b) $I(X;Y) := H(X) - H(X \mid Y)$
   c) $I(X;Y) := H(X \mid Y) - H(X)$
   d) $I(X;Y) := H(Y \mid X) - H(Y)$

4. What is the entropy of a uniform distribution on a set of n elements?

   a) $\frac{n}{\log n}$
   b) n
   c) $\log n$
   d) $n \log n$

5. What is the definition of conditional expectation?

   a) $\Pr(A \mid B) := \frac{\Pr(A \wedge B)}{\Pr(B)}$
   b) $\Pr(A \mid B) := \Pr(A \wedge B) \times \Pr(A)$
   c) $\Pr(A \mid B) := \frac{\Pr(A \wedge B)}{\Pr(A)}$

d) $\Pr(A \mid B) := \Pr(A \wedge B) \times \Pr(A)$

6. One error correction strategy is to send the message repeatedly, as long as it has not been received correctly by the recipients. This strategy raises a number of issues. Which of the following statements is *wrong* and *not* an inherent problem of this strategy?
   a) The latency between message emission and reception in some applications is too high for this scheme to be feasible.
   b) It is impossible for a recipient to determine with high probability whether the message received is correct.
   c) For some use cases (e.g., storage for later retrieval), it may not be possible to resend the message.
   d) The strategy may require too many transmission rounds to be feasible if the error probability is too high.

7. Consider information sent over a binary symmetric channel with error probability f. An R_3 repetition code (that repeats each bit three times) is used as an error correcting code. What is the remaining bit error probability?
   a) $1-(1-f)^3$
   b) f^3
   c) $f+3f^2(1-f)+f^3$
   d) $3f^2(1-f)+f^3$

8. The proof of Shannon's theorem uses a trick, what does it consist in?
   a) If the average error probability of a code, averaged over all codes, is greater than α, then there exists a code with average error probability greater than α.
   b) If the average error probability of a code, averaged over all codes, is less than α, then there exists a code with average error probability less than α.
   c) If the average error probability of a code, averaged over all codes, is less than α, then there exists no code with average error probability greater than α.
   d) If the average error probability of a code, averaged over all codes, is greater than α, then there exists no code with average error probability less than α.

9. Which of the following equalities is always true for arbitrary random variables X, Y, Z?
   a) $I(X;YZ) = I(Y;Z \mid X)$
   b) $I(X;YZ) = I(X;Y) + I(X;Y \mid Z)$
   c) $I(X;YZ) = I(X;Y) + I(X;Z \mid Y)$
   d) $I(X;YZ) = I(X;Y \mid Z) + I(X;Z \mid Y)$

10. Which of the following coding scheme uses the same amount of bits to encode each input block:
    a) Hamming error correction
    b) Arithmetic coding
    c) Huffman coding
    d) Lempel-Ziv compression

11. Let $L(C,X)$ be the expected codelength of an optimal symbol code for random variable X. Which of the following inequalities is always correct? *Because of an error in the proposed choices, there are actually two correct answers, choose either.*
    a) $H(X) \leqslant L(C,X) \leqslant H(X)+1$
    b) $H(X)-1 \leqslant L(C,X) \leqslant H(X)$
    c) $L(C,X) \leqslant H(X) \leqslant L(C,X)+1$
    d) $L(C,X)-1 \leqslant H(X) \leqslant L(C,X)$

12. The capacity of a channel is the mutual information:
    a) averaged over all source probability distributions
    b) potentially achieved by a number of source probability distributions
    c) achieved for a single source probability distribution
    d) never achieved but the limit of mutual information achieved by all source probability distributions

13. Consider the following codeword lengths: 1, 3, 3, 4, 4, 5, 6, 6. Which of the following is correct:
    a) There is a unique uniquely decodable complete code with these code lengths.
    b) There are several uniquely decodable complete codes with these code lengths.
    c) There is no uniquely decodable symbol code with these code lengths.
    d) There is an incomplete uniquely decodable code with these code lengths.

14. Consider a binary source X emitting i.i.d. bits. We define the β-typical set as

$$T_{N\beta} := \left\{ x^N \in \{0,1\}^N : \left| \frac{1}{N} \log \frac{1}{\Pr(X^N = x^n)} - H(X) \right| < \beta \right\}$$

    Which of the following statement is correct?
    a) For N large enough, $\Pr(X \in T_{N\beta})$ is arbitrarily low
    b) For N large enough, $\Pr(X \in T_{N\beta})$ is arbitrarily close to $H(X)$
    c) For N large enough, $\Pr(X \in T_{N\beta})$ is arbitrarily close to $\frac{1}{2}$
    d) For N large enough, $\Pr(X \in T_{N\beta})$ is arbitrarily high

15. The weak law of large numbers is of the form: $\Pr(|X - \alpha| \geqslant \beta) \leqslant \frac{\gamma^2}{\beta^2 N}$. What are α, β, and γ in this formula?
    a) α is the expected value of X, β the standard deviation, γ an arbitrary constant
    b) α is the expected value of X, β an arbitrary constant, γ the standard deviation
    c) α is an arbitrary constant, β the standard deviation of X, γ the expected value
    d) α is an arbitrary constant, β the expected value of X, γ the standard deviation

16. What is the rate of the $(7,4)$ Hamming code?
    a) $7/3$
    b) $7/4$
    c) $4/7$
    d) $3/7$

17. In the $(7,4)$ Hamming code, error-correcting bits are defined as follows (addition is modulo 2): $t_5 := s_1 + s_2 + s_3$, $t_6 := s_2 + s_3 + s_4$, $t_7 := s_1 + s_3 + s_4$. What is the most likely decoding of 1010100?
    a) 1010
    b) 1110
    c) 1011
    d) 0010

18. Which of the following statements is *incorrect*:
    a) There is always an optimal uniquely decodable code.
    b) All prefix codes are uniquely decodable.
    c) There is always an optimal prefix code.
    d) All uniquely decodable codes are prefix codes.

19. Shannon's source coding theorem for lossy compression schemes tells us:
    a) that as long as the information loss probability is higher than a threshold, we can compress to any rate, if the message length is high enough.
    b) that it is possible to compress a message with arbitrarily low probability of information loss as close as desired from the entropy, as long as the message length is high enough.
    c) that a long enough message can be compressed without any loss of information at a rate that is equal to the entropy.
    d) that the entropy defines the limit of how much we can compress the message, independently of the probability of information loss.

20. Which of the following statement about error correction is true:
    a) For every channel, for every rate, arbitrary low error probability can be achieved provided the block length is high enough.
    b) For every channel, for every block length, arbitrary low error probability can be achieved provided the rate is low enough.
    c) For every channel, there is a minimum error probability that can be achieved by any code.
    d) For every channel, arbitrary low error probability can be achieved provided the rate is low enough and the block length is high enough.

21. Which of the following statement is *incorrect*?
    a) Arithmetic coding encodes a full message as a fraction of the $[0, 1]$ interval.
    b) Lempel-Ziv is a data compression scheme that relies on a dictionary of prefixes encountered so far.
    c) Huffman coding is an optimal symbol code for lossless data compression.
    d) Stream codes, contrarily to symbol codes with small block length, can encode messages in a streaming fashion.

22. Let X, Y be two random variables with integer values. In which of the following situations can we have $H(X \mid Y) \neq 0$?
    a) Y has the same set of values as X
    b) X is such that there exists an x with $\Pr(X = x) = 1$
    c) $X = Y^2$ (the value of X is the square of the value of Y)
    d) $X = Y$

23. What is the capacity of the binary symmetric channel with error probability f?
    a) $1 + f \log f + (1 - f) \log(1 - f)$ bits
    b) $1 + f \log f$ bits
    c) $-f \log f$ bits
    d) $-f \log f - (1 - f) \log(1 - f)$ bits

24. Consider an ensemble with probabilities $\{\frac{1}{2}, \frac{1}{4}, \frac{3}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{64}\}$. Construct an Huffman code for this ensemble. What is the sum of the lengths of all codewords?
    a) 18
    b) 24
    c) 22
    d) 20

25. Consider an ensemble with probabilities $\{\frac{1}{2}, \frac{1}{4}, \frac{3}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{64}\}$ (as in the previous question). Construct an Huffman code for this ensemble. What is the expected codeword length?
    a) $\frac{37}{16}$ bits
    b) $\frac{69}{32}$ bits

c) $\frac{59}{32}$ bits
d) 2 bits

26. If $X \to Y \to Z$ is a Markov chain of random variables, we always have:
    a) $I(X;Y) \leqslant I(X;Z)$
    b) $I(X;Y) \geqslant I(X;Z)$
    c) $I(X;Z) = I(X;Y) + I(Y;Z)$
    d) $I(X;Z) = I(X;Y) + H(Y \mid Z)$

27. What is the Shannon entropy of a random variable X?
    a) $H(X) = \sum_x \Pr(X = x) \log \frac{1}{\Pr(X=x)}$
    b) $H(X) = \sum_x \frac{1}{\Pr(X=x)} \log \Pr(X = x)$
    c) $H(X) = \sum_x \Pr(X = x) \log \Pr(X = x)$
    d) $H(X) = \sum_x \frac{1}{\Pr(X=x)} \log \frac{1}{\Pr(X=x)}$

28. Assume X and Y are two independent random variables. Which of the following equations does not necessarily hold?
    a) $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$
    b) $H(X,Y) = H(X) + H(Y)$
    c) $I(X;Y) = 0$
    d) $H(X \mid Y) = H(Y \mid X)$

29. If we try sending a message at a rate higher than the capacity of a channel:
    a) We cannot achieve arbitrarily low bit error probability.
    b) We can achieve arbitrarily low bit error probability, but only for arbitrarily long messages.
    c) We necessarily have bit error probability approaching 1 as the message length increases.
    d) It is impossible to use a channel at a rate higher than its capacity.

30. A bomb detector, that has probability 10^{-6} of incorrectly detecting a bomb when there is none (and always detects a bomb when there is one), is installed at Changi airport. One estimates that only 1 person out of 100 million may carry a bomb. What is the probability that a passenger actually carries a bomb if the detector detects one?
    a) $\approx 10^{-3}$
    b) $\approx 10^{-6}$
    c) $\approx 10^{-2}$
    d) $\approx 1 - 10^{-6}$

END OF PAPER