THE STATE UNIVERSITY OF NEW JERSEY
RUTGERS

College of Engineering
Department of Electrical and Computer Engineering

330:542 **Information Theory and Coding** **Spring 1996**

Quiz 2

You have 80 minutes to answer the following 3 questions. Please read the exam in its entirety before beginning work. Please underline and/or box your final answers for each part to aid in the grading process. You are allowed two sides of an 8.5X11 standard sheet of note paper, a pencil and a blue book. YOU MUST SHOW ALL WORK TO RECEIVE CREDIT.

The problems may be hard or easy. I can never tell ahead of time.

GOOD LUCK!

1. *(30 points)* **Channel Capacity Snippets:**

   (a) *(10 points)* You are given N cascaded binary symmetric channels each with crossover probability ϵ. Assuming no processing between stages, what is the capacity of the N cascaded channels? What distribution achieves capacity?

   (b) *(10 points)* Must a "stongly" symmetric (i.e., not weakly symmetric) channel be square? Square means the channel has equal numbers of inputs and outputs. Prove if true/ give counterexample if false.

   (c) *(10 points)* Prove or disprove: *The overall channel formed by the cascade of any N k-input/k-output symmetric channels is itself a symmetric channel.*

2. *(30 points)* **Coding Theorem:** In our proof of the channel coding theorem, we assumed that the W messages coded by the X^n were equally likely. Start from the assumption that you have $W = 2^{nR}$ messages and explain why rate parameter R implies that the W should be equally likely.

   HINT: What if the messages were not equally likely? Think of the source-coding theorem.

3. *(40 points)* **More Channel Capacity:**

   It is often useful to be able to calculate channel capacity. Searching over input distributions $p(\mathbf{x})$ to find the maximum mutual information is what we'll explore in this problem.

   NOTE: I composed this problem but have not done it yet, so it may be hard (or easy).

   Fly-By-Night, an optimization corporation tells you that they have a procedure based on gradient descent which finds the optimal input distribution using the first partials you calculated in the first part of this problem. Their procedure is as follows:

- Make an initial guess $\mathbf{p}(0)$
- Calculate the gradient $\nabla I(\mathbf{p}, \mathbf{Q})$ with respect to the input probability vector $\mathbf{p}$, where $I(\mathbf{p}, bfQ)$ is the mutual information function and $\mathbf{Q}$ is the transition probability matrix.
- Iterate

$$\mathbf{p}(n+1) = \nabla I(\mathbf{p}(n), \mathbf{Q})\delta + \mathbf{p}(n)$$

  where δ is a suitably small number.

(a) *(10 points)* Please justify the use of the gradient to find the unique absolute maximum of the mututal information function.

(b) *(10 points)* Try the FBN procedure on a binary symmetric channel with crossover probability ϵ. Does it work? Why/why not?

(c) *(20 points)* If you answered in the previous part that "it works" and you're correct, then you can skip this part and pick up the 20 points essentially for free. However, if you answered that the procedure does not work, please modify the proposed gradient procedure appropriately and show that your procedure works on the binary symmetric channel with crossover probability ϵ and the binary erasure channel with erasure probability δ regardless of the initial guess $\mathbf{p}(0)$.