MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science
Department of Mechanical Engineering

6.050J/2.110J **Information and Entropy** **Spring 2004**

Issued: March 8, 2004 **Problem Set 6** **Due: March 12, 2004**

Problem 1: Prime Massachusetts Madness

In Massachusetts primary elections, there are two ballots, one for Republicans (R) and one for Democrats (D) (in this problem, ignore the other, small parties). Voters can be registered with either of the two parties or can be independent (I). Independent voters may, on the day of the primary, enroll in either party before voting (and then may revert to independent status immediately after voting if they wish).

Ideally, all registered Republicans (R) would get a Republican ballot (RB) and all Democrats (D) would receive a Democrat ballot (DB). However, one year there were numerous complaints and you decide to use Bayes' Theorem to analyze what went on. You are told that in two different cities, Newton and Cambridge, the poll workers did not hand out the ballots properly. In particular, 20% of the registered Republicans were given Democrat ballots by mistake, and 30% of the registered Democrats received Republican ballots. Half of the independent voters received Republican ballots and half Democrat ballots.

a. What are the desired conditional probabilities $p(RB \mid R)$, $p(RB \mid D)$, $p(DB \mid R)$, and $p(DB \mid D)$? What were the actual conditional probabilities? What were the actual probabilities $p(RB \mid I)$ and $p(DB \mid I)$?

b. An analysis of the voter registration list (which is a public document) showed that in Newton 30% of those registered were Democrats and 20% Republicans. Your friend Alice, from Newton, was seen with a Republican ballot. What are the probabilities that she was a registered Democrat (D), a registered Republican (R), or an independent (I)? In other words, find $p(D \mid RB)$, $p(R \mid RB)$, and $p(I \mid RB)$.

c. In Cambridge, the voters are registered 60% Democrat and 20% Republican, with the rest independent. Bob, who lives in Cambridge, received a Republican ballot. If you had to guess whether he was registered Republican, Democrat, or Independent, which is most likely? What are the probabilities of those three $p(D \mid RB)$, $p(R \mid RB)$, and $p(I \mid RB)$?

Problem 2: Special Orders Don't Upset Us

Buzz, the hot new dining spot on campus, emphasizes simplicity. It only has two items on the menu, burgers and zucchini. Customers make a choice as they enter (they are not allowed to order both), and inform the cooks in the back room by shouting out either "B" or "Z". Unfortunately the two letters sound similar so 10% of the time the cooks misinterpret what was said. The marketing experts who designed the restaurant guess that 95% of the orders will be for burgers and 5% for zucchini.

The cooks can hear one order per second. The customers arrive at the rate of one every second. One of the chefs says that this system will never work because customers send orders at exactly the rate at which they can be accepted, so you could barely keep up even if there were no noise in the channel. You are hired as an outside consultant to deal with the problem.

a. What is the channel capacity C of this communication channel in bits per second?

b. What is the information content per order?

c. Is it possible, through some coding arrangement, to transmit orders reliably at your target rate of one per second?

d. If the restaurant becomes more popular, what is the maximum rate your system can handle, measured in orders per second (assuming you could find an effective source code and channel code)?

e. You decide to process orders two at a time. Devise a Huffman code in which you transmit the information for two orders in less than 2 bits on average. Will this scheme work, given the channel capacity?

Problem 3: Groceries a Dodo

The new internet-savvy grocery store chain, the Digital Deli, has recently conceived of a new way to generate business. The company plans to sell internet-enabled refrigerators, brand name Smart-and-Cool. Sensors in the appliance keep track of the supply of various items, such as milk, beer, or ketchup and when the amount on hand gets low, an order is sent via the internet to the Digital Deli, whereupon the item is delivered to the customer the next day. This problem will deal with designing the protocol which allows the refrigerators to communicate with the store computer, making use of IP with its tendency to drop packets when the network is overloaded.

The company first had their marketing manager, Debbie Dodo, design the protocol. The protocol she designed was as follows: First, the Smart-and-Cool decides what products it needs to order. Then it sends a start-of-order packet (STO) to the Digital Deli central server. The server responds with a ready packet (RDY). Then the refrigerator sends a number of item packets (ITM) which contain the name of an item and the quantity requested. It sends these serially, one after the other. Finally, it sends an end-of-order packet (EDO) and the Digital Deli server responds with an invoice packet (INV) which details the total cost of the order and the expected time of delivery.

a. There are several major problems with Debbie's protocol. Specifically, during the alpha testing, customers complained that some items ordered were not delivered. Describe in words a scenario in which IP packet dropping might produce this result. Also, the store manager noticed that sometimes the server was waiting indefinitely for the next ITM packet. Explain how this could happen.

When the managers realized that Debbie Dodo's protocol was seriously flawed, they hired an engineer, Ben Bitdiddle, to do a redesign. Ben added two features which he though would improve the protocol: receipt packets (RCT) and timeouts. RCT packets are sent by the server to a refrigerator whenever an ITM packet is received. Timeouts are used by the refrigerator to determine when to resend a packet when an expected response is not received. Under this scheme the refrigerator resends the STO packet if after a certain amount of time (the *timeout*) it does not receive the RDY packet, resends an ITM packet if within that same amount of time it does not receive the corresponding RCT packet, and resends the EDO packet if it does not receive the INV packet.

b. There are problems with Ben's design. Specifically, consider the following occurrence: a customer's refrigerator decides that it is low on ketchup and places an order for one bottle to the Digital Deli. The next day the customer receives a shipment of 10,000 bottles of ketchup. What went wrong? In addition to explaining this problem, write down at least one more problem that Ben's design would encounter.

Losing patience, the managers of the Digital Deli decide to bite the bullet and hire the best-of-the-best, you, an MIT graduate and engineer.

c. Redesign the protocol to solve the problems enumerated in the first two parts above.

Turning in Your Solutions

If you used MATLAB for this problem set, you may have some M-files and a diary. You may turn in this problem set by e-mailing your M-files and diary to 6.050-submit@mit.edu. Do this either by attaching them to the e-mail as *text* files, or by pasting their content directly into the body of the e-mail (if you do the latter, please indicate clearly where each file begins and ends). Alternatively, you may turn in your solutions on paper in room 38-344. The deadline for submission is the same no matter which option you choose.

Your solutions are due 5:00 PM on Friday, March 12, 2004. Later that day, solutions will be posted on the course website.