CSC 2508S
Spring, 2002

September 18, 2002
DUE: October 2, 2002

Information: Quantification, Representation and Manipulation

Assignment #1

1. [30 marks]

Consider a game in which one player flips a fair coin while the other rolls a fair die. Denote the outcome of the coin by $C \varepsilon \{H,T\}$, and the outcome of the die by $D \varepsilon \{1,2,3,4,5,6\}$. Let the "scores" of the two players be denoted by S_1 and S_2, respectively.

Define the scores obtained by the two players by:

$$S_1 = \begin{cases} 6 & \text{if } C=H \\ 1 & \text{if } C=T \end{cases} \qquad \text{and} \qquad S_2 = D$$

and the "winning" variable, W, to be:

$$W = \begin{cases} +1 & \text{if } S_1 > S_2 \\ 0 & \text{if } S_1 = S_2 \\ -1 & \text{if } S_1 < S_2 \end{cases}$$

Evaluate each of the following quantities using the "facts" presented in lecture (which make many of the parts trivial to calculate). Indicate the basis for each answer.

i.	$H(D)$	ix.	$H(W,C)$
ii.	$H(C)$	x.	$I(D,C)$
iii.	$H(W)$	xi.	$I(D,W)$
iv.	$H(W \mid D)$	xii.	$I(C,W)$
v.	$H(W \mid C)$	xiii.	$H(W \mid C,D)$
vi.	$H(D \mid W)$	xiv.	$H(C \mid W,D)$
vii.	$H(C \mid W)$	xv.	$H(D \mid W,C)$
viii.	$H(W,D)$		

b. For the fifteen parts of a, how do the answers change if each "C" is replaced by "S_1", and each "D" is replaced by "S_2"? Explain your answer.

c. How do the answers to part a change if the definition of S_1 is changed to be:

$$S_1 = \begin{cases} 5 & \text{if } C=H \\ 2 & \text{if } C=T \end{cases}$$

2. [30 marks]
The populations of twelve cities in Canada's provinces and territories are as follows:

City	*Population*	*City*	*Population*
Edmonton, ALB	917,500	Yellowknife, NWT	37,360
Vancouver, BC	1,995,900	Toronto, ONT	4,594,900
Winnepeg, MAN	676,400	Charlottetown, PEI	32,531
St. John, NB	127,300	Montreal, PQ	3,428,300
St. John's, NF	173,600	Saskatoon, SASK	229,300
Halifax, NS	348,000	Whitehorse, YUK	31,070
Pond Inlet, NUN	1,154		

a) How much information is conveyed by learning the city of residence of a randomly selected person who lives in one of the 12 cities?

b) How much information is conveyed by learning whether or not a randomly selected person lives in Montreal?

c) How much information is conveyed by learning whether or not a randomly selected person lives in the Yellowknife?

d) Determine a *binary* Huffman Code for encoding the city of residence of each member of a large group of people who live in one of the 12 cities.

e) Determine a *ternary* Huffman code for the same situation.

f) Determine a *decimal* Huffman code for the same situation.

g) Calculate and compare the relative redundancy of the three codes in parts (d), (e), and (f).

3. [40 marks]
Pick a full-information board game with which you are familiar (such as checkers, chess or go) where two players alternate moves, and there are no elements of chance. Consider encoding various *positions* that arise during a game. The purpose of such an encoding might be to store very concisely a record of all positions encountered in a series of tournaments, where the total number of games to be recorded is huge.

A. Given a huge number of "positions" drawn from actual games (i.e., resulting from a sequence of legal moves from the initial position), specify an encoding for a position with as little redundancy as you can manage. About how many bits would be required to encode a position on average?

B. Briefly comment on the related problem of encoding a complete "game" (a sequence of positions starting with the initial position, and related by legal moves).

2. [20 marks]
The populations of Canada's provinces and territories are as follows:

Province	*Population*	*Province*	*Population*
Alberta	2,545,553	Ontario	10,084,885
British Columbia	3,282,061	Prince Edward Island	129,765
Manitoba	1,091,942	Quebec	6,895,963
New Brunswick	723,900	Saskatchewan	998,928
Newfoundland	568,474	Northwest Territories	57,600
Nova Scotia	899,942	Yukon	27,797

A.

i. How much information is conveyed by learning the province or territory of residence of a randomly selected Canadian resident?

ii. How much information is conveyed by learning whether or not a randomly selected Canadian resident lives in Ontario?

iii. How much information is conveyed by learning whether or not a randomly selected Canadian resident lives in the Yukon?

B.

i. Determine a *binary* Huffman Code for encoding the Province or territory of residence of each member of a large group of Canadian residents.

ii. Determine a *ternary* Huffman code for the same situation.

iii. Determine a *decimal* Huffman code for the same situation.

C. Calculate and compare the *relative redundancy* of the three codes in part B.

Edmonton, ALB 917,500 Toronto, ONT 4,594,900 Vancouver, BC 1,995,900 Charlottetown, PEI 32,531 Winnepeg, MAN 676,400 Montreal, PQ 3,428,300 St. John, NB 127,300 Saskatoon, SASK 229,300 St. John's, NF 173,600 Yellowknife, NWT 37,360 Halifax, NS 348,000 Whitehorse, YUK 31,070 Nunavut, NUN 26,745