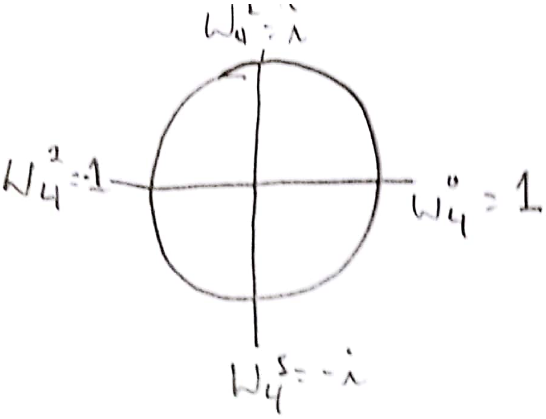
Assignment 5

Ryan Chalmers

Question 1

- We know from class that the entropy provides an absolute limit on the best lossless encoding of any message, assuming the message may be represented as a sequence independent identically distributed random variables.
- However, we know that the message is not independent, since each word must be separated by a space, the number of spaces is directly dependent on the number of words.
- We will attempt to produce an encoding that is better than the entropy by using Huffman. We know Huffman is within 0.586 of the entropy. In an attempt to improve upon this we will take into account that every word in the message is from the english language. If we preprocess the words in the dictionary we can can use the fact that in the English language some letters are more likely to be proceeded by others (e.g. t is likely proceeded by h). Therefore we can give each letter in the alphabet its own Huffman encoding that uses the increased probability of the letters proceeding it to improve upon Huffman. By doing this we are using the extra information given by the question to generate an encoding that is better than the entropy.

Question 2



$$\begin{bmatrix} (W_4^0)^0 & (W_4^0)^1 & (W_4^0)^2 & (W_4^0)^3 \\ (W_4^1)^0 & (W_4^1)^1 & (W_4^1)^2 & (W_4^1)^3 \\ (W_4^2)^0 & (W_4^2)^1 & (W_4^2)^2 & (W_4^2)^3 \\ (W_4^3)^0 & (W_4^3)^1 & (W_4^3)^2 & (W_4^3)^3 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ -1-i \\ 0 \\ -1+i \end{bmatrix}$$

Question 3

$$p1 = a(f - h) \qquad p2 = (a + b)h$$
$$p3 = (c + d)e \qquad p4 = d(g - e)$$
$$p5 = (a + d)(e + h) \qquad p6 = (b - d)(g + h)$$
$$p7 = (a - c)(e + f)$$

The A x B can be calculated using above seven multiplications.
Following are values of four sub-matrices of result C

$$\underbrace{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}_{A} \times \underbrace{\begin{bmatrix} e & f \\ g & h \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} p5 + p4 - p2 + p6 & p1 + p2 \\ p3 + p4 & p1 + p5 - p3 - p7 \end{bmatrix}}_{C}$$

- To be able to solve the an instance of the 16x16 matrix multiplication using the Strassen algorithm we must recursively divide each matrix into 4 sub-matrices each of size n/2 by n/2, until we reach matrices of size 2x2.
- In other words to multiply one 2x2 we use 7 multiplications, a 4x4 uses 7^2 multiplications, a 8x8 uses 7^3 multiplications, and a 16x16 uses 7^4.

Question 4

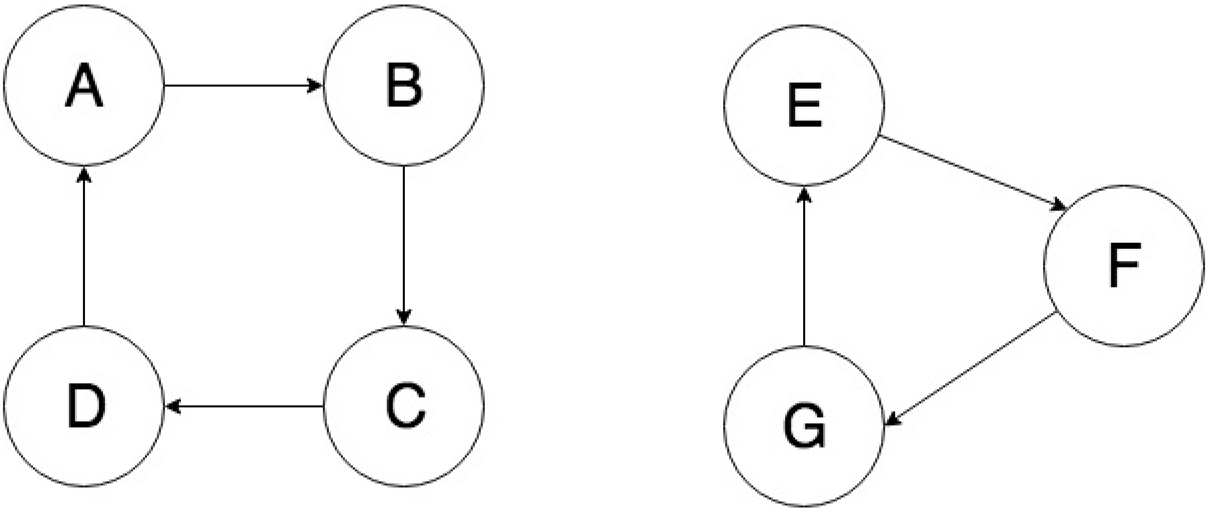
- Suppose we have some instance of Huffman encoding where there are two symbols, one has a very high probability (approaching 1) and one has a very low probability (approaching 0).
- We can represent these two symbols using one bit, therefore R(S) = 1
- According to the formula given for entropy in class:

$$= -\sum P(X_i)\log_b P(X_i)$$

- Calculating H(S) using our chosen probabilities we get a entropy H(S) which is near 0.
- In this case R(S) ≤ H(S) + 0.586 does not hold.

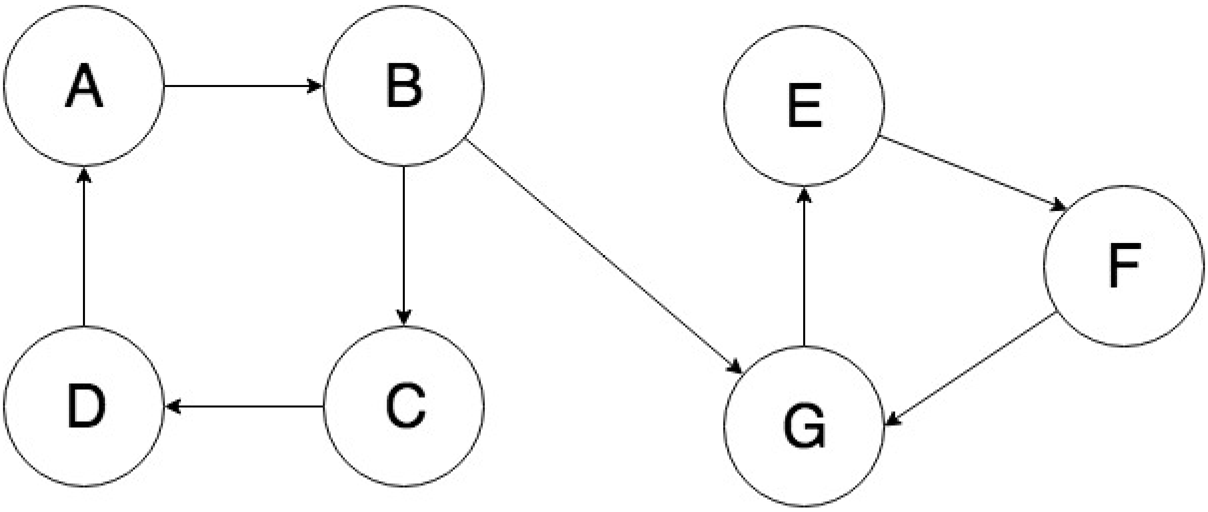
Question 5

5.a.



- In this case since we have 7 vertices and each are in their own cycle, each vertex will rank 1/7 and have an equal page rank of 1.

5.b.



- In order to find the page rank we attempt to solve (P-I)R = 0.

$$\begin{bmatrix} \frac{0.15}{7}-1 & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7}+0.85 & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7} \\ \frac{0.15}{7}+0.85 & \frac{0.15}{7}-1 & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7} \\ \frac{0.15}{7} & \frac{0.15}{7}+\frac{0.85}{2} & \frac{0.15}{7}-1 & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7} \\ \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7}+0.85 & \frac{0.15}{7}-1 & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7} \\ \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7}-1 & \frac{0.15}{7} & \frac{0.15}{7}+0.85 \\ \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7}+0.85 & \frac{0.15-1}{7} & \frac{0.15}{7} \\ \frac{0.15}{7} & \frac{0.15}{7}+\frac{0.85}{2} & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7} & \frac{0.15}{7}+0.85 & \frac{0.15}{7}-1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

- Solved using gaussian elimination.

x_1	0.34133
x_2	0.37777
x_3	0.24813
x_4	0.29851
x_5	0.93756
x_6	0.88455
x_7	1.00000

- Now we can rank the vertices (1 being the highest, 7 the lowest).

Rank	1	2	3	4	5	6	7
Vertex	G	E	F	B	A	D	C