You are encouraged to collaborate on the homework and ask for assistance. You are required to write your own solutions, list your collaborators, acknowledge all sources of help, and cite all external references. Failure to follow these guidelines will be considered a breach of academic honesty regulations.

Submit your solution by Wednesday October 8 in class or electronically via the link on the course webpage. Late submissions won't be accepted.

Question 1

Let S be the symmetric matrix

$$\begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix}.$$

(a) Run power iteration on S with initialization $\mathbf{a} = (1, 0, 0)$. What is the state $\mathbf{a}_5$ after five steps? What is the spectral norm estimate $\alpha_5 = \|S\mathbf{a}_5\|/\|\mathbf{a}_5\|$?

(b) Repeat part (a) with initialization $\mathbf{b} = (0, 1, 0)$. How do $\mathbf{b}_5$ and β_5 compare with $\mathbf{a}_5$ and α_5 above? Explain the similarities and differences.

(c) Now repeat parts (a) and (b) on input $S' = S + I$, where I is the identity matrix. Explain how the answers in this part relate to the ones you obtained in parts (a) and (b).

(d) Finally, pick any vector $\mathbf{d}$ orthogonal to the output $\mathbf{c}_5$ from part (c) (when initialized with $(1, 0, 0)$). What happens when you run power iteration on S' with initialization $\mathbf{d}$? Explain.

Question 2

Let $f(x, y) = \phi(x - y) + \phi(x + y - 1)$, where ϕ is some "activation function" with one input and one output.

(a) Draw a circuit for f with $+$, $\times$, and ϕ gates. The inputs should be x, y, and the constant 1 and edge weights should be used for scaling.

(b) Draw the circuit ∇f obtained by applying Backpropagation to f. The gates in your circuits should be $+$, $\times$, ϕ, and ϕ'. What is the size of this circuit? What is the depth?

(c) Assume now that ϕ is the function $\phi(t) = \frac{1}{2}t^2$. Then $\phi'(t) = t$ is the identity function. Simplify the circuit by evaluating all gates that depend on constants only (e.g., replace $1 \times (-3)$ by the constant -3) and short-circuiting all ϕ' gates. What is the size and depth of the simplified circuit?

(d) Explain how the circuit in part (c) computes the function $\nabla f(\mathbf{x}) = A^T(A\mathbf{x} + b)$, where $A\mathbf{x} = b$ is the linear system $x - y = 0, x + y = 1$.

(e) In general, given a linear system with m equations and n unknowns, what is the size, depth, and number of wires of the output of backpropagation given a circuit for $f(\mathbf{x}) = \frac{1}{2}\|A\mathbf{x} - b\|^2$ as input? Provide an answer in big-theta notation and justify it.

Question 3

A *decision tree* is a nested if-then-else program that branches over variables and outputs variables or their negations. For example, the decision tree in Figure 1 evaluates to $+1$ when the input is $x = +1, y = +1, z = +1$.

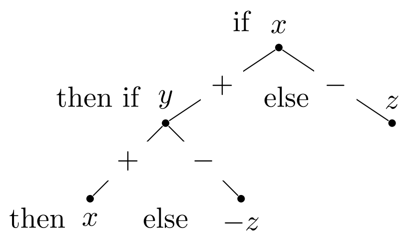

Figure 1: A function $f(x, y, z)$ specified by a decision tree.

(a) Write the list representation of the function f specified by the decision tree in Figure 1.

(b) Write the polynomial representation of f. (**Hint:** For every root-to-leaf path write a polynomial that outputs the leaf value if this path is taken and zero if not.)

(c) The *depth* of the decision tree is the maximum number of nested loops plus one, which equals the length of the longest root-to-leaf path in the tree representation plus one. For example the decision tree in Figure 1 has depth three. Prove that for any decision tree, the degree of its polynomial representation is at most its depth.

(d) The *2-bit addressing function* is the function $\mathrm{addr}_2\colon \{-1, +1\}^6 \to \{-1, +1\}$ given by

$$\mathrm{addr}_2(x, y, z_{++}, z_{+-}, z_{-+}, z_{--}) = z_{xy}.$$

Describe a decision tree for addr_2. What is the degree of its polynomial representation? Prove both an upper bound (it is at most ...) and a lower bound (it is at least ... because ...).

Question 4

You have recorded the graph of mutual friendships among sixteen people (Alice, Bob, up to Patrick). They cluster into two groups. Most of the friendships are among people within the same group. Find the two groups. Your personalized instance is available here: `https://andrejb.net/csi4103/hw/25H02.html`.

Your solution should consist of two lists, one for the members in each group (e.g., group 1: Alice, Charlie, Fred; group 2: Bob, Dave, Eve).

You may use any of the algorithms covered in class or write your own. Explain clearly how you arrived at your solution. Undocumented computer code will not be entertained as a satisfactory explanation.