Caio Brighenti
COSC 480 - Learning From Data
Fall 2019
Problem Set 6

1. Problem 2.3

   (a) Positive or negative ray

   Given a hypothesis set consisting of either a positive or negative ray, the total number of dichotomies will be equal to $|D_1 \cup D_2|$, where D_1 is the dichotomy set for positive rays, and D_2 is the dichotomy set for negative rays. As defined earlier in the chapter, the number of dichotomies for a ray is $N+1$. Thus, we have that $|D_1 \cup D_2| = (N+1)+(N+1)-|D_1 \cap D_2|$. The set $D_1 \cap D_2$ clearly will only have two elements: the dichotomy when h for all values in X is $+1$, and when h is -1 for all $x \in X$.
   Therefore, the total number of dichotomies is $m_h(N) = |D_1 \cup D_2| = (N+1)+(N+1)-2 = 2N$.

   (b) Positive or negative interval

   We approach this problem in the same manner as the one before. The total number of dichotomies is equal to $|D_1 \cup D_2| = |D_1| + |D_2| - |D_1 \cap D_2|$, where D_1 is the set dichotomies for positive intervals and D_2 is the set of dichotomies for negative intervals. In both the negative and positive case, the maximum number of dichotomies is $m_h(N) = \frac{1}{2}N^2 + \frac{1}{2}N + 1$. It remains to be shown what the set $D_1 \cap D_2$ is composed of.
   The only cases where a positive and negative interval overlap is when at least one endpoint of the interval is at x_1 or x_n. In any of these cases, the intervals that produce the same dichotomy set across positive and negative intervals can be expressed by $[x_1, a]$ and $[a, x_n]$, where one of these is the endpoints of a positive interval and the other a negative interval. Thus we must find the number of possible intervals with at least one endpoint equal to x_1 or x_n.
   This is clearly the same as the case above, where we select an endpoint a and the second endpoint from either x_1 or x_n. Thus, $|D_1 \cap D_2| = 2N$, and therefore $m_h(N) = |D_1 \cup D_2| = \frac{1}{2}N^2 + \frac{1}{2}N + 1 + \frac{1}{2}N^2 + \frac{1}{2}N + 1 - 2N = N^2 + N + 2 - 2N = N^2 - N + 2$.

   (c) Two cocentric spheres

   While this problem appears highly complicated at first, it is simply a specialized version of the interval problem. Each dichotomy is expressed as the radius of a sphere, cocentric with the spheres of radius a and b. A radius is $+1$ if it falls between the range a, b, and -1 otherwise. This is clearly just an interval on the radii produced by X, and thus the maximum number of dichotomies is simply $\frac{1}{2}N^2 + \frac{1}{2}N + 1$.