CNS/CS/EE 188a Computation Theory and Neural Systems - Homework #5

Ling Li, ling@cs.caltech.edu

February 20, 2001

5.1 Weight Size and Boolean Symbols. I assume the mapping from $\{0,1\}$ to $\{-1,1\}$ is $0 \mapsto -1$, $1 \mapsto 1$.

(a) $f(x_1, x_2, x_3) = \text{sgn}(-4 + x_1 + 2x_2 + 5x_3)$ for $x_i \in \{0,1\}$. $S(w) = 4 + 1 + 2 + 5 = 12$. It is easy to see that when $x_3 = 0$, $f = 0$ and when $x_3 = 1$, $f = 1$. Thus f can be written as $f = \text{sgn}(x_3 - 1)$, with $S(w') = 2 < S(w)$.

(b) $\hat{S}(OR) = n + 1$ over $\{0,1\}$. For any w_i implementing OR, consider inputs below:

- $x_1 = x_2 = \cdots = x_n = 0$. $OR = \text{sgn}(w_0) = 0$, so $w_0 \le -1$.
- $x_i = 1$, $x_j = 0\ \forall j \ne i$. $OR = \text{sgn}(w_0 + w_i) = 1$, so $w_0 + w_i \ge 0$, or $w_i \ge -w_0 \ge 1$.

Thus $\sum_{i=0}^n |w_i| \ge n + 1$. Since for

$$OR(x_1, x_2, \ldots, x_n) = \text{sgn}\left(-1 + x_1 + x_2 + \cdots + x_n\right) \tag{1}$$

we have $\sum_{i=0}^n |w_i| = n + 1$, we know $\hat{S}(OR) = n + 1$ over $\{0,1\}$.

(c) $\hat{S}(OR) = 2n - 2$ over $\{-1,1\}$. Let $W = \sum_{i=1}^n w_i$. Consider two cases of inputs in (b) with 0 replaced by -1. That is

$$\begin{aligned} w_0 - W &\le -1, & (2)\\ w_0 + 2w_i - W &\ge 0. & (3) \end{aligned}$$

Substituting (3) by (2), we have $2w_i \ge 1$, or $w_i \ge 1$ since all the weights are integers. Thus $W \ge n$. Summing (3) over $i = 1, \ldots, n$, we have

$$nw_0 + 2W - nW \ge 0.$$

That is, $w_0 \ge \frac{n-2}{n} W \ge n - 2$. So $\sum_{i=0}^n |w_i| \ge 2n - 2$.
We can transform (1) which is over $\{0,1\}$ to $OR = \text{sgn}\left(n - 2 + x_1 + \cdots + x_n\right)$ over $\{-1,1\}$. For this representation, we have $\sum_{i=0}^n |w_i| = 2n - 2$. So $\hat{S}(OR) = 2n - 2$ over $\{-1,1\}$ (for $n \ge 2$).

(d) $MAJ(x_1, x_2, \ldots, x_n) = \text{sgn}\left(-\frac{n+1}{2} + x_1 + \cdots + x_n\right)$ over $\{0,1\}$.

Proof: Note that n is odd, so $w_0 = -\frac{n+1}{2}$ is an integer. $\sum_{i=1}^n x_i$ is the number of 1's in the inputs. If more than half of the inputs are 1, $\sum_{i=1}^n x_i \ge \frac{n+1}{2}$ and $MAJ = 1$; If not more than half of the inputs are 1, that is, $\sum_{i=1}^n x_i \le \frac{n-1}{2}$, we have $w_0 + \sum_{i=1}^n x_i \le -1$ and $MAJ = 0$. Therefore the majority function is a threshold function.

(e) $\hat{S}(MAJ) = \frac{3n+1}{2}$ over $\{0,1\}$. For any w_i implementing the majority function, consider an arbitrary subset I of $\{1,2,\dots,n\}$ with size k. If the inputs are $x_i = 1$ iff $i \in I$, then $MAJ = 1$ iff $k \geq \frac{n+1}{2}$. Let's look at a subset I with size $\frac{n+1}{2}$:

- When the inputs are $x_i = 1$ iff $i \in I$, we get $MAJ = 1$ and

$$w_0 + \sum_{i \in I} w_i \geq 0. \tag{4}$$

- When the inputs are $x_i = 1$ iff $i \in I - \{j\}$, where j is an arbitrary element in I, we have $MAJ = 0$ and

$$w_0 + \sum_{i \in I} w_i - w_j = w_0 + \sum_{i \in I - \{j\}} w_i \leq -1. \tag{5}$$

From above two inequalities, we get $w_j \geq 1$. Since I and j can both be arbitrary, we know $w_i \geq 1 \forall i \geq 1$. Then from (5),

$$w_0 \leq -1 - \sum_{i \in I - \{j\}} w_i \leq -1 - (\frac{n+1}{2} - 1) = -\frac{n+1}{2}.$$

So for any weights w_i that implements MAJ, we get $\sum_{i=0}^{n} |w_i| \geq \frac{n+1}{2} + n = \frac{3n+1}{2}$. However, the threshold function in (d) has the weights sum $\frac{3n+1}{2}$. Thus we know $\hat{S}(MAJ) = \frac{3n+1}{2}$ over $\{0,1\}$.

(f) $\hat{S}(MAJ) = n$ over $\{-1,1\}$. First, we can find such threshold function doing MAJ with weights sum n:

$$MAJ(x_1, x_2, \dots, x_n) = \operatorname{sgn}(x_1 + \cdots + x_n).$$

The proof that above function is MAJ is straightforward: more than half of the inputs are 1 $\Leftrightarrow$ there are more 1's than -1's $\Leftrightarrow$ the above function returns 1. Now we want to prove this is the minimum weights sum.

The two inequalities in (e) now become

$$\begin{aligned} w_0 - \sum_{i=1}^{n} w_i + 2\sum_{i \in I} w_i &\geq 0, \\ w_0 - \sum_{i=1}^{n} w_i + 2\sum_{i \in I - \{j\}} w_i &\leq -1. \end{aligned}$$

Thus we have $2w_j \geq 1$, or $w_j \geq 1$ since all weights are integers. So we also get $w_i \geq 1 \forall i \geq 1$. Thus $\sum_{i=0}^{n} |w_i| \geq \sum_{i=1}^{n} |w_i| \geq n$. So, $\hat{S}(MAJ) = n$ over $\{-1,1\}$.

(g) Since $n + 1 < 2n - 2$ for $n \geq 4$, $\hat{S}(OR)$ over $\{0,1\}$ is smaller than $\hat{S}(OR)$ over $\{-1,1\}$, for $n \geq 4$. However, since $\frac{3n+1}{2} > n$, $\hat{S}(MAJ)$ over $\{0,1\}$ is larger than $\hat{S}(MAJ)$ over $\{-1,1\}$, for all n. Thus for an arbitrary threshold function f, $\hat{S}(f)$ over $\{0,1\}$ is not always smaller than $\hat{S}(f)$ over $\{-1,1\}$.

5.2 Batcher Sorting Network.

(a) The depth, by induction, is $\log_2 n$, where n should be some power of 2.

Proof: Merge[2] is a single comparator, so the depth is 1, i.e., $\log_2 2$. Suppose the depth of Merge[n] is $\log_2 n$. Since the Merge[$2n$] network consists of two Merge[n] networks and 1 more layer of comparators, the depth of Merge[$2n$] is $\log_2 n + 1$, i.e., $\log_2 2n$.

(b) The depth, by induction, is $\frac{1}{2}\log_2 n(\log_2 n + 1)$ (thus is $O(\log_2^2 n)$).

Proof: Batcher[2] is a single comparator, so the depth is 1. And $\frac{1}{2}\log_2 2(\log_2 2 + 1)$ is also 1. Suppose the depth of Batcher[n] is $\frac{1}{2}\log_2 n(\log_2 n+1)$. Since the Batcher[$2n$] network consists of two Batcher[n] and 1 Merge[$2n$], the depth of Batcher[$2n$] is

$$\frac{1}{2}\log_2 n(\log_2 n + 1) + \log_2 2n = \frac{1}{2}(\log_2 n + 2)(\log_2 n + 1) = \frac{1}{2}\log_2 2n(\log_2 2n + 1).$$

(c) A sorting network with four inputs consists of 5 comparators is shown in Figure 1. The

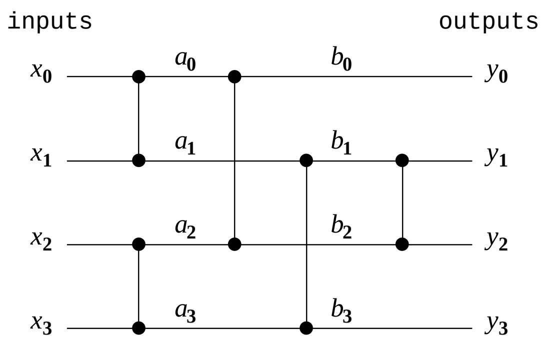

Figure 1: Sorting network with 5 comparators.

idea is to find the minimum and maximum of the four inputs in the first 2 stages and then sort the middle 2 in the third stage. Since $a_0 = \min\{x_0, x_1\}$ and $a_2 = \min\{x_2, x_3\}$, we get $y_0 = b_0 = \min\{a_0, a_2\} = \min\{x_0, x_1, x_2, x_3\}$. Similarly, $y_3 = \max\{x_0, x_1, x_2, x_3\}$. Since only comparators are used, no numbers are lost — so y_1 and y_2 are two median values of the inputs, and they are sorted. So this is a sorting network.

5.3 Smoothing Networks.

(a) The outputs are $(y_0, y_1, y_2, y_3) = (4, 4, 5, 5)$ (Figure 2).

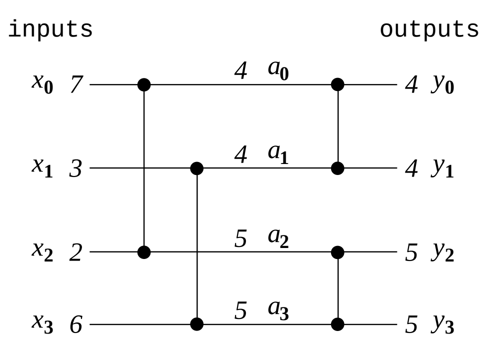

Figure 2: the Smooth[4] network.

(b) The property of the balancer gives $a_0 \leq a_2$, $a_1 \leq a_3$, and $y_0 \leq y_1$, $y_2 \leq y_3$ in Figure 2. Thus

$$\begin{aligned} y_0 = \left\lfloor \frac{a_0 + a_1}{2} \right\rfloor &\leq \left\lfloor \frac{a_2 + a_3}{2} \right\rfloor = y_2, \\ y_1 = \left\lceil \frac{a_0 + a_1}{2} \right\rceil &\leq \left\lceil \frac{a_2 + a_3}{2} \right\rceil = y_3. \end{aligned}$$

So y_0 is no larger than all other y_i and is the smallest output; y_3 is no smaller than all other y_i and is the largest output. (d) gives the maximum difference between y_0 and y_3 is 2.

(c) The Smooth[8] network is shown in Figure 3. Please see (d) for the proof that the maximum

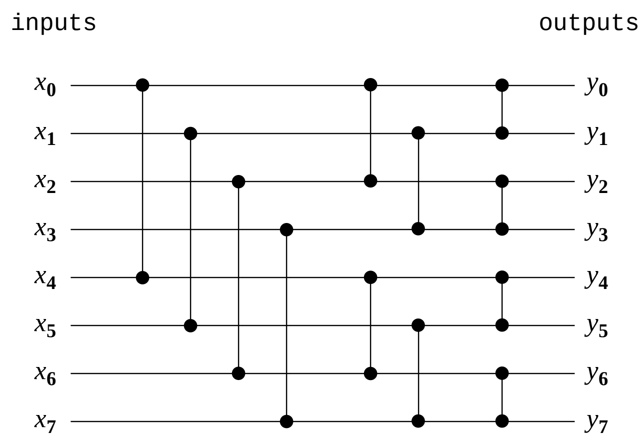

Figure 3: the Smooth[8] network.

difference between y_0 and y_7 is 3.

(d) For the Smooth[n] network, we have below claim:

Claim: The outputs of the Smooth[n] network is within the range $[\bar{x} - \frac{1}{2}\log n, \bar{x} + \frac{1}{2}\log n]$, where $n \geq 2$ and $\bar{x}$ is the mean of the inputs. And there exist inputs so that $y_0 = \bar{x} - \frac{1}{2}\log n$ and $y_{n-1} = \bar{x} + \frac{1}{2}\log n$.

Proof: For $n = 2$, the Smooth[n] is a single balancer and the outputs for inputs x_0, x_1 are $\lfloor \bar{x} \rfloor$ and $\lceil \bar{x} \rceil$, where $\bar{x} = \frac{x_0 + x_1}{2}$ is the mean of the inputs. Since the inputs are integers, $\bar{x}$ is also some integer, or some interger plus $1/2$. In either case the outputs are within $[\bar{x} - \frac{1}{2}, \bar{x} + \frac{1}{2}]$. For inputs $x_0 = 0$ and $x_1 = 1$, $\bar{x} = \frac{1}{2}$, $y_0 = 0$ and $y_1 = 1$, i.e., $y_0 = \bar{x} - \frac{1}{2}\log n$ and $y_1 = \bar{x} + \frac{1}{2}\log n$. So the claim holds for $n = 2$.

Suppose the claim holds for some $n \geq 2$. The Smooth[$2n$] network (Figure 4) consists of n balancers in the first stage and two Smooth[n] networks thereafter. Let a_i denotes the outputs of the first stage. There exist inputs $z_0, z_1, \ldots, z_{n-1}$ so that for the Smooth[n] network, $y_0 = \bar{z} - \frac{1}{2}\log n$ and $y_{n-1} = \bar{z} + \frac{1}{2}\log n$. Let $x_i = z_i$ and $x_{n+i} = z_i + 1$ for $0 \leq i < n$. Thus for these inputs, $\bar{x} = \bar{z} + \frac{1}{2}$, and $a_i = x_i$ for all i. Thus the first Smooth[$2n$] network gives

$$y_0 = \bar{z} - \frac{1}{2}\log n = \bar{x} - \frac{1}{2}\log 2n.$$

And the second Smooth[$2n$] network gives (here I skip the simple proof that for a Smooth network, if we add 1 to all the inputs, then we will get all the outputs increased

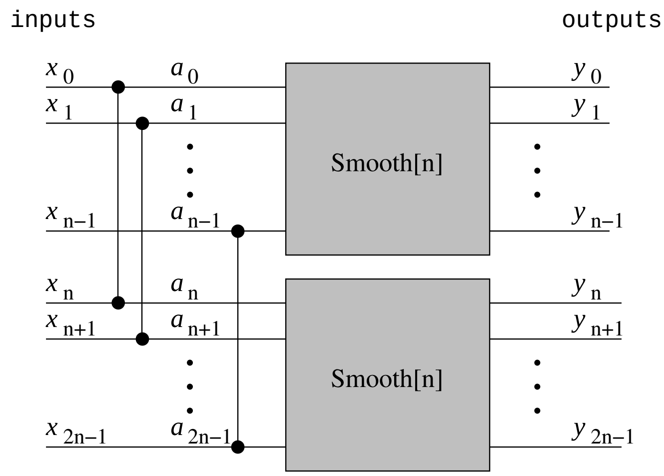

Figure 4: Constructing the $\texttt{Smooth}[2n]$ network from $\texttt{Smooth}[n]$ networks.

by 1. In general, if we perform a linear transformation $ax+b$ where a, b are both integers to all the inputs, then we will get all the outputs also transformed by $ax+b$. This is similar to what a monotonic increasing function does to a sorting network.)

$$y_{2n-1} = \bar{z} + 1 + \frac{1}{2}\log n = \bar{x} + \frac{1}{2}\log 2n.$$

For all balancers ($\texttt{Smooth}[2]$) in the first stage, we have

$$a_i, a_{n+i} \in \left[\frac{x_i + x_{n+i}}{2} - \frac{1}{2}, \frac{x_i + x_{n+i}}{2} + \frac{1}{2}\right], \tag{6}$$

for $0 \le i < n$, and for those two $\texttt{Smooth}[n]$, we have

$$\begin{aligned} y_i &\in \left[\frac{1}{n}\sum_{j=0}^{n-1} a_j - \frac{\log n}{2}, \frac{1}{n}\sum_{j=0}^{n-1} a_j + \frac{\log n}{2}\right], \\ y_{n+i} &\in \left[\frac{1}{n}\sum_{j=0}^{n-1} a_{n+j} - \frac{\log n}{2}, \frac{1}{n}\sum_{j=0}^{n-1} a_{n+j} + \frac{\log n}{2}\right]. \end{aligned}$$

(6) gives

$$\frac{1}{2n}\sum_{j=0}^{2n-1} x_j - \frac{1}{2} \le \frac{1}{n}\sum_{j=0}^{n-1} a_j, \frac{1}{n}\sum_{j=0}^{n-1} a_{n+j} \le \frac{1}{2n}\sum_{j=0}^{2n-1} x_j + \frac{1}{2}.$$

Therefore for all i,

$$y_i \in \left[\frac{1}{2n}\sum_{j=0}^{2n-1} x_j - \frac{\log n + 1}{2}, \frac{1}{2n}\sum_{j=0}^{2n-1} x_j + \frac{\log n + 1}{2}\right],$$

which is just $y_i \in \left[\bar{x} - \frac{1}{2}\log 2n, \bar{x} + \frac{1}{2}\log 2n\right]$. So the claim also holds for $2n$. $\square$

Thus, the maximum difference between y_0 and y_{n-1} is $\log n$, which is also achievable.