Q520: Answers to the Homework on Hopfield Networks

1. For each of the following, answer true or false with an explanation:
   
   a. Fix a Hopfield net. If $\vec{o}$ and $\vec{o}'$ are neighboring observation patterns then $\Phi(\vec{o})$ and $\Phi(\vec{o}')$ cannot be equal. This is true for our example, but false in general. For example, consider the case when each $w_{ij}$ is zero. In this case, all the energies of all possible input vectors are 0 as well. So for all $\vec{o}$ and all neighbors $\vec{o}'$ of it, $\Phi(\vec{o}) - \Phi(\vec{o}') = 0$.

   b. The synchronous algorithm for Hopfield nets never goes into an infinite loop. This is false as we pointed out in an example in class.

   c. The sequential algorithm for Hopfield nets never goes into an infinite loop. This is true as explained in class. The sequential algorithm has the property that the small steps always either leave the current vector alone, or move to a neighbor of strictly decreasing energy. The whole net is finite, so the energy cannot decrease forever. Thus there must be local minima of energy. (Actually, it is a general fact that if we have any function at all from a finite graph to the natural numbers or the real numbers, then there must be local minima, and if we wander around from neighbor to neighbor always going strictly down in energy, we must reach one of those local minima.)

   d. If we run a the sequential algorithm starting with one of the target patterns, we always get the very same pattern back. This is false, though we would like it to be true! The main example from class showed this (again): two of the target patterns turned out not to be local minima of energy. For another example, take the case $n = 2$, and try to store $(1, 1), (-1, 1), \text{ and } (1, -1)$. The weight $w_{1,2}$ turns out to be $-1$. The energy function is $\Phi(1, 1) = 1, \Phi(1, -1) = -1, \Phi(-1, 1) = -1, \text{ and } \Phi(-1, -1) = -1$. In particular, note that $(1, 1)$ is not a local minimum of energy. So if we run the net starting with it, we would move to $(1, -1)$ or $(-1, 1)$.

2. Let’s consider the case of a Hopfield net for the two patterns $x^1 = (1, 1, 1, 1)$ and $x^2 = (-1, -1, -1, -1)$.

   a. Find the weights of the net. For $i \neq j$, $w_{i,j} = (1)(1) + (-1)(-1) = 2$.

   b. Show that the stored patterns $x^1$ and $x^2$ are local minima of the energy function $\Phi(\vec{o})$. The formula in this case for energy is

   \[
   \Phi(o_1, o_2, o_3, o_4) = -\frac{1}{2} \sum_{i \neq j} w_{i,j} o_i o_j = -\sum_{i \neq j} o_i o_j
   \]

   Note now that $\Phi(x^1) = \sum_{i \neq j} 1 = -12$, and similarly for $\Phi(x^2)$. Where do we get $-12$? There are 12 terms in the energy formula, and for the input vectors $x^1$ and $x^2$ the values of $o_i o_j$ are always 1. We also have a negative sign outside.
And for any other vector $\vec{o}$ besides $x^1$ and $P(-1)$, the sum $\sum_{i \neq j} o_i o_j$ is going to contain some negative number. So it will be smaller than $-12$. Thus $\Phi(\vec{o})$ will be bigger than $-12$ in this case. So we have shown that $x^1$ and $x^2$ are absolute minima of $\Phi$ in this net. Hence they are local minima as well.

c. Conversely, show that if $\vec{o}$ is a local minimum of $\Phi(\vec{o})$, then $\vec{o}$ must be either $x^1$ or $x^2$. So with this $x^1$ and $x^2$, the net always converges to one of the patterns that we have stored in it. Probably the easiest way here is by brute force. Here is a table of all the energies:

<table>
<thead>
<tr>
<th>$\Phi(1, 1, 1, 1)$</th>
<th>$-12$</th>
<th>$\Phi(-1, 1, 1, 1)$</th>
<th>$0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Phi(1, 1, 1, -1)$</td>
<td>$0$</td>
<td>$\Phi(-1, 1, 1, -1)$</td>
<td>$4$</td>
</tr>
<tr>
<td>$\Phi(1, 1, -1, 1)$</td>
<td>$4$</td>
<td>$\Phi(-1, 1, -1, 1)$</td>
<td>$4$</td>
</tr>
<tr>
<td>$\Phi(1, -1, 1, -1)$</td>
<td>$4$</td>
<td>$\Phi(-1, -1, 1, -1)$</td>
<td>$0$</td>
</tr>
<tr>
<td>$\Phi(1, -1, -1, 1)$</td>
<td>$4$</td>
<td>$\Phi(-1, -1, -1, 1)$</td>
<td>$0$</td>
</tr>
<tr>
<td>$\Phi(1, -1, -1, -1)$</td>
<td>$0$</td>
<td>$\Phi(-1, -1, -1, -1)$</td>
<td>$-12$</td>
</tr>
</tbody>
</table>

As you can see, every point whose energy is 4 has a neighbor with energy 0, and every point whose energy is 0 has a neighbor with energy $-12$. The only local minima occur at the points of energy $-12$.

d. Suppose we run the sequential algorithm starting on $(1, 1, -1, -1)$. Let’s say that we pick our update vertices randomly. Show that the output pattern can be either $x^1$ or $x^2$, and that the choice depends entirely on which update vertex we pick first. If we decide to update vertex 1 first, we would re-set it to $-1$. And then, when we got to 2, we would also update it to $-1$. So in the end, we get $(-1, -1, -1, -1) = x^2$. But if we decided to update vertex 3 first, we would make it 1. Then when 4 gets updated, the same thing happens. So we get $(1, 1, 1, 1) = x^1$.

3. Let’s say that a list of input vectors $x^1, x^2, \ldots, x^n$ is storable if the set of local minima of the Hopfield net for these vectors is exactly the vectors on this list.

Is the list $(1, 1, 1, 1), (-1, -1, -1, -1), (1, 1, -1, -1), (-1, -1, 1, 1)$ storable? Yes. The weights are $w_{12} = w_{34} = 4$, and all others are 0. Here is a table of energies:

<table>
<thead>
<tr>
<th>$\Phi(1, 1, 1, 1)$</th>
<th>$-8$</th>
<th>$\Phi(-1, 1, 1, 1)$</th>
<th>$0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Phi(1, 1, 1, -1)$</td>
<td>$0$</td>
<td>$\Phi(-1, 1, 1, -1)$</td>
<td>$8$</td>
</tr>
<tr>
<td>$\Phi(1, 1, -1, 1)$</td>
<td>$0$</td>
<td>$\Phi(-1, 1, -1, 1)$</td>
<td>$8$</td>
</tr>
<tr>
<td>$\Phi(1, -1, 1, -1)$</td>
<td>$-8$</td>
<td>$\Phi(-1, 1, -1, -1)$</td>
<td>$0$</td>
</tr>
<tr>
<td>$\Phi(1, -1, -1, 1)$</td>
<td>$0$</td>
<td>$\Phi(-1, -1, -1, 1)$</td>
<td>$0$</td>
</tr>
<tr>
<td>$\Phi(1, -1, -1, -1)$</td>
<td>$8$</td>
<td>$\Phi(-1, -1, -1, -1)$</td>
<td>$-8$</td>
</tr>
</tbody>
</table>

So the target list gives absolute minima and hence local minima.
4. If $\vec{o} = -\vec{o}'$, then $\Phi(\vec{o}) = \Phi(\vec{o}')$.

[a.] Why is this the case?

$$\Phi(\vec{o}) = -\frac{1}{2} \sum_{i,j} w_{i,j} o_i o_j = -\frac{1}{2} \sum_{i,j} w_{i,j} (-o_i)(-o_j) = -\frac{1}{2} \sum_{i,j} w_{i,j} o'_i o'_j = \Phi(\vec{o}').$$

[b.] Use part (a) to show that if $\vec{o}$ is a local minimum of energy, so is $\vec{o}'$. The point here is that the negatives of neighbors of $\vec{o}'$ are all neighbors of of $\vec{o}$. To see this, take some neighbor of $\vec{o}'$. For some $i$, our neighbor of $\vec{o}'$ is

$$(-o_1, \ldots, -o_{i-1}, o_i, -o_{i+1}, \ldots, -o_n). \tag{1}$$

The negative of this is

$$(o_1, \ldots, o_{i-1}, -o_i, o_{i+1}, \ldots, o_n).$$

This differs from $\vec{o}$ in exactly one place, so it is a neighbor of $\vec{o}$.

Now we assume that $\vec{o}$ is a local minimum of energy, and we prove the same thing about $\vec{o}'$. Take any neighbor of $\vec{o}'$. For example, take the one in (1). We want to show that

$$\Phi(\vec{o}') \leq \Phi(-o_1, \ldots, -o_{i-1}, o_i, -o_{i+1}, \ldots, -o_n). \tag{2}$$

As we know, the negative of the vector in (1) is a neighbor of $\vec{o}$. So since $\vec{o}$ is local minimum of energy, we have

$$\Phi(\vec{o}) \leq \Phi(o_1, \ldots, o_{i-1}, -o_i, o_{i+1}, \ldots, o_n). \tag{3}$$

But by part (a), $\Phi(\vec{o}) = \Phi(\vec{o}')$, and again by part (a),

$$\Phi(-o_1, \ldots, -o_{i-1}, o_i, -o_{i+1}, \ldots, -o_n) = \Phi(o_1, \ldots, o_{i-1}, -o_i, o_{i+1}, \ldots, o_n). \tag{4}$$

(This last step is really crucial.) So putting (3) together with (4) and $\Phi(\vec{o}) = \Phi(\vec{o}')$ gives (2). And this is our goal.

5. The problem on Hopfield nets with bias values on the nodes.

[a.] Let $\vec{o}$ and $\vec{o}'$ be neighbors, agreeing at all entries except the $i$th. Then

$$\Phi(\vec{o}') - \Phi(\vec{o}) = 2 \theta_i o_i + \sum_j o_j w_{i,j}.$$
we’ll say $\Phi(\vec{o}) = -\sum_j \theta_j o_j - \frac{1}{2} \sum_{j,k} w_{j,k} o_j o_k$. Then

$$
\Phi(\vec{o}') - \Phi(\vec{o})
= - \left( \sum_j \theta_j o_j' - \frac{1}{2} \sum_{j,k} w_{j,k} o_j' o_k' \right) - \left( - \sum_j \theta_j o_j - \frac{1}{2} \sum_{j,k} w_{j,k} o_j o_k \right)
= - \left( \sum_j \theta_j o_j' + \sum_j \theta_j o_j \right) - \frac{1}{2} \sum_{j,k} \left( w_{j,k} o_j' o_k' - w_{j,k} o_j o_k \right)
= - (\theta_i o_i' - \theta_i o_i) - \sum_k \left( w_{i,k} o_i' o_k' - w_{i,k} o_i o_k \right)
= - (\theta_i o_i - \theta_i o_i) - \sum_k w_{i,k} \left( -o_i o_k - o_i o_k \right)
= 2\theta_i o_i + 2o_i \sum_k w_{i,k} o_k
= 2o_i(\theta_i + \sum_j o_j w_{i,j})
$$

The most important line to be sure about is the one marked (*). The point is that in the sum $\sum_{j,k} w_{j,k} o_j' o_k' - w_{j,k} o_j o_k$, for each $j$ and $k$ that are both different from $i$, the differences cancel. So in the sum as a whole, we only have to consider the differences $w_{j,k} o_j' o_k' - w_{j,k} o_j o_k$ when $j$ or $k$ (or both) are $i$. But when $j = k$, $w_{j,k} = 0$. Anyway, the differences with $j = i$ and other $k$ are exactly the same as the differences with $k = i$ and other $j$. This is because $w_{j,k} o_j o_k = w_{k,j} o_k o_j$. So we are left with a double listing of the differences involving $i$. That is why the sum in the line below (*) has only one variable, and also why the $\frac{1}{2}$ has gone away.

[b.] Use part (a) to prove the main fact about these kinds of Hopfield nets:

**Lemma** Let $\vec{o}$ be the vector at some stage of the algorithm. If a small step of the algorithm changes $\vec{o}$ to $\vec{o}'$, then we must have $\Phi(\vec{o}') < \Phi(\vec{o})$. And if some neighbor of $\vec{o}$ has smaller energy than $\vec{o}$, then some small step of the algorithm will change $\vec{o}$.

Here is the proof: First, consider the small step for vertex $i$. Let $\vec{o}'$ be just like $\vec{o}$ except that $o_i' = -o_i$. Suppose that the small step changes $o_i$ to $o_i'$. Then the reason that we did this is because $(\theta_i + \sum_j o_j w_{i,j})$ is not zero, and indeed it has opposite sign from $o_i$. (Otherwise, we would not change $o_i$ at all.) It follows that

$$
2o_i(\theta_i + \sum_j o_j w_{i,j}) < 0. \tag{5}
$$

And then by part (a), $\Phi(\vec{o}') - \Phi(\vec{o}) < 0$. This means that $\Phi(\vec{o}') < \Phi(\vec{o})$.

In the other direction, suppose that neighbor of $\vec{o}$ has smaller energy than $\vec{o}$. Let $\vec{o}'$ be the first that the algorithm comes to that are of smaller energy than $\vec{o}'$. We show that the some small step for $i$ will change $\vec{o}$. The reason is similar to the paragraph above, only sdrawkcb. Suppose that $\Phi(\vec{o}') < \Phi(\vec{o})$. Then it follows that (5) holds. So $o_i$ and $\theta_i + \sum_j o_j w_{i,j}$ have opposite signs (and the latter is not zero). Then by the way we update the net, we definitely change $o_i$ to $-o_i$ in this small step. That is, we change $\vec{o}$ to $\vec{o}'$. 

[4]
6. Find a Hopfield net with bias values on the nodes which has \( x^1 = (1, 1, 1, 1) \) as its only storable vector. Before we set the \( \theta \) values, we calculate the weights as earlier. We get \( w_{i,j} = 1 \) all \( i \neq j \). Here are the energy values:

\[
\begin{align*}
\Phi(1, 1, 1, 1) &= -6 & \Phi(-1, 1, 1, 1) &= 0 \\
\Phi(1, 1, 1, -1) &= 0 & \Phi(-1, 1, 1, -1) &= 2 \\
\Phi(1, 1, -1, 1) &= 0 & \Phi(-1, 1, -1, 1) &= 2 \\
\Phi(1, -1, -1, -1) &= 2 & \Phi(-1, -1, -1, -1) &= 0 \\
\Phi(1, 1, -1, -1) &= 2 & \Phi(-1, -1, 1, 1) &= 0 \\
\Phi(1, -1, -1, 1) &= 2 & \Phi(-1, 1, -1, 1) &= 0 \\
\Phi(1, -1, -1, -1) &= 0 & \Phi(-1, 1, -1, -1) &= -6 \\
\end{align*}
\]

We have two energy minima, at \((1,1,1,1)\) and \((-1,-1,-1,-1)\). To eliminate the second of these, we add numbers to each entry on behalf of each \(-1\), and subtract on behalf of each \(1\). We have to do some guessing to see how much to do. But after some trial and error, we settle on \( \theta_i = 4 \) for \( i = 1, 2, 3, 4 \). When we do this, we get

\[
\begin{align*}
\Phi(1, 1, 1, 1) &= -6 - 16 = -22 & \Phi(-1, 1, 1, 1) &= 0 - 8 = -8 \\
\Phi(1, 1, 1, -1) &= 0 - 8 = -8 & \Phi(-1, 1, 1, -1) &= 2 \\
\Phi(1, 1, -1, 1) &= 0 - 8 = -8 & \Phi(-1, 1, -1, 1) &= 2 \\
\Phi(1, -1, -1, -1) &= 2 & \Phi(-1, -1, -1, -1) &= 0 + 8 = 8 \\
\Phi(1, 1, -1, 1) &= 0 - 8 = -8 & \Phi(-1, -1, 1, 1) &= 2 \\
\Phi(1, -1, -1, 1) &= 2 & \Phi(-1, -1, -1, 1) &= 0 + 8 = 8 \\
\Phi(1, -1, -1, -1) &= 0 + 8 = 8 & \Phi(-1, -1, -1, -1) &= -6 + 16 = 10 \\
\end{align*}
\]

Let’s check that each point other than \((1,1,1,1)\) has a neighbor with strictly smaller energy. Flip any \(-1\) to a \(1\), and observe that the energy decreases.

Summary: to get a Hopfield net which successfully stores \((1,1,1,1)\) and nothing else, we take \( w_{i,j} = 1 \) for \( i \neq j \), and also \( \theta_i = 4 \) for \( i = 1, \ldots, 4 \).