Minds and Computers

• Discovering the nature of intelligence by studying intelligence in all its forms: human and machine

• Artificial intelligence (A.I.)
  – The science of getting computers to behave in an intelligent manner
  – Useful for applications of computers to diagnosis, speech recognition, chess playing, handwriting recognition, credit checks, search engines, etc.
  – Want computers to perform as well as possible

• Computational simulation of human cognition
  – The science of getting computers to behave in human-like ways
  – Want computers to make the same mistakes that people do
Why build computer simulations of human mental processes?

- **Learn from mother nature**
  - Take advantage of millions of years of “research and development” into human minds by evolution

- **Force psychologists to be precise**
  - Computers only do what told in a precise language

- **Insure that the theory will work**
  - Need computer to keep track of parts of complicated theories

- **Obtain quantitative predictions about behavior**

- **Two different kinds of computational simulation**
  - Explicit knowledge-engineering
  - Neural Networks (Connectionism)
Explicit Knowledge Representation

• Programmer feeds in all of the information the computer needs to know
• “If…Then” rules that determine computer’s actions
• ELIZA as a simulation of a psychological therapist
  – “IF person mentions ‘father’ THEN say ‘Who else in your family comes to mind when you think about this?’
  – “IF person writes ‘I remember X’ THEN say ‘Does it make you feel happy to recall X?’
  – ELIZA has no intelligence itself. Intelligence comes from people interpreting its statements.
Neural Networks (Connectionism)

- **Inspired by brains**
  - Simple neuron-like units, massively interconnected
  - Parallel processing

- **Units**
  - Activation = Activity of unit
  - Weights = Strength of the connection between two units
    - Excitatory and inhibitory

- **Advantages**
  - Learning = changing strength of connection between units
  - Graceful degradation = if part of the system is lesioned, then system still works. Performance slowly degrades as more of the system is lesioned.
  - Noise tolerance = system still works if noise is added to inputs
  - Content addressable memory = memories are not searched like items in a list. The most appropriate memory is directly accessed.
Hebbian Learning

- “Units that fire together, wire together”

\[ W_{i,j} = L S_i S_j \]

\[ W_{1,3} = 0.4 \]

\[ W_{1,2} = -0.4 \]

\[ W_{2,3} = -0.4 \]

\[ S_1 = 1 \]
\[ S_2 = -1 \]
\[ S_3 = 1 \]
\[ Prediction_Y = \sum_{X=1}^{N} S_X W_{X,Y} \]

\[ Prediction_3 = 1 \times 0.4 + (-1) \times (-0.4) = 0.8 \]
One unit per cell
If White, $S=-1$
If Black, $S=+1$

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$-1$</td>
<td>$+1$</td>
<td>$-1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$-1$</td>
<td>$+1$</td>
<td>$+1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$-1$</td>
<td>$-1$</td>
<td>$+1$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Positive Weights

Weights to here

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$-1$</td>
<td>$+1$</td>
<td>$-1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$-1$</td>
<td>$+1$</td>
<td>$+1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$-1$</td>
<td>$-1$</td>
<td>$+1$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

States

Weights to here
Connections between units change when “A” is presented. Connection strength increases if the units are in the same state, and decreases otherwise.

Missing information for incomplete pattern is “filled in.” Cells that are off are turned on because they have positive connections to other cells that are on.