# Exercise: Trace Backpropagation

Consider a neural net with one hidden layer, two inputs *a* and *b*, one hidden unit *c*, and one output unit *d*. The activation function is the sigmoid for each node. This network has five weights (*wac, wbc, w0c, wcd, w0d)*, where *w0x* represents the bias or threshold weight for unit *x*. Initialize these weights to the values (.1,.1,.1,.1,.1), then give their values after each of the first two training iterations of Backpropagation algorithm. Assume learning rate (step size) of 0.3, stochastic (incremental) gradient descent (without momentum), and the following training examples:

|  |  |  |  |
| --- | --- | --- | --- |
| Data Point | a | b | d |
| x1 | 1 | 0 | 1 |
| x2 | 0 | 1 | 0 |

1. Using the notation in the slides, show the formulas for computing the following quantities, for any node *x:*
   * ax
   * Δ[x]
   * Weight update forwxy
2. Fill in the following table using the formulas from the slides. You can expand this to include more information (e.g. derivatives of activation functions) if you like.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Data Point | ac | Δ[c] | ad | Δ[d] | w0c | wac | wbc | wcd | w0d |
| x1 |  |  |  |  |  |  |  |  |  |
| x2 |  |  |  |  |  |  |  |  |  |