

HW 2 (PROBLEM 8)

If the transfer function of the hidden units is linear, then a three-layer network is equivalent to two layer as any combination of linear functions always gives another linear function.

e.g. consider 2 linear functions:-

$$f_1(x) = w_1 x_1 + w_2 x_2 + b_1 = y_1$$

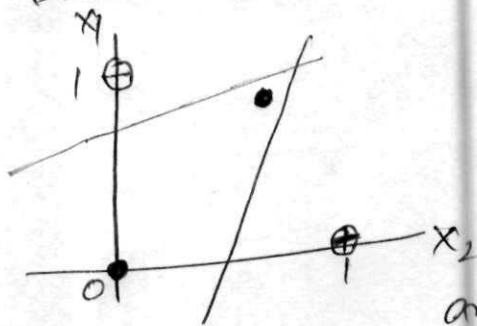
$$f_2(x) = w_3 x_1 + w_4 x_2 + b_2 = y_2$$

If we combine these two functions using another linear function e.g.

$$g(x) = w_5 y_1 + w_6 y_2 + b_3$$

This is another linear function of original inputs. Hence a 3 layer network with hidden ~~units~~ units with a linear function is equivalent to a 2 layer network.

XOR can't be solved with such a network as it is not a linearly separable problem. XOR function can be plotted as shown below:-



There is no combination of lines (linear functions) which can separate \oplus & \bullet points.

Same argument can be extended to any n-bit parity function whose o/p is 1 only if there are odd numbers of 1.