HW1 R06922063

1.

(1) NO, 3-2-2-1 network "can not" implement all 2^{2^3} functions, 因 input data 在第一個 2个 neuron 的 hidden layer 就一定會把某些 function (例如 XOR) 中不同 class 的資料切到同一個 region 之中, 造成 1st hidden layer 的 output 出的 binary coding 一定有 "ambiguous representation", 根據第4章講義, upper layer 永遠無法正確分類出已經有 ambiguous representation 的 lower layer output, 故 3-2-2-1 network 也無法分出如 XOR、XNOR 等 function, 因第二層 hidden layer 已經不可能把第一層 hidden layer 分錯的 data 校正。

二个 hyperplane 無法切出三維的 XOR

(2) N-N-1 network 表示我們在 hidden layer 中有 N 個 N-1 維的 hyperplane、XOR 定義為在 input 中有奇數個維度的值為 1 時輸出 1，反之輸出 0、對 N 維 input $(x_1, x_2, \dots x_n)$, hidden layer 的 N 个 hyperplane 的方程式可以設成: (設 activation 為大於 0 變成 1，小於 0 變成 0)

$$① \quad x_1 + x_2 + x_3 \dots + x_n = 1 - \varepsilon$$
$$② \quad \frac{x_1}{2} + \frac{x_2}{2} + \frac{x_3}{2} \dots + \frac{x_n}{2} = 1 - \varepsilon$$
$$③ \quad \frac{x_1}{3} + \frac{x_2}{3} + \frac{x_3}{3} \dots + \frac{x_n}{3} = 1 - \varepsilon$$
$$\vdots$$
$$ⓝ \quad \frac{x_1}{n} + \frac{x_2}{n} + \frac{x_3}{n} \dots + \frac{x_n}{n} = 1 - \varepsilon$$

(ε 為一小於 $\frac{1}{n}$ 的數)

→ 第ⓧ个 hyperplane 只有在 data 有 x 以上个維度為 1 時會輸出 1

這 n 個 hyperplane 總共只會生出 n 種有效的 code: (0, 0, 0, …, 0), (1, 0, 0 … 0) (1, 1, 0, … 0) (1, 1, 1, … 0) (1, 1, 1, …, 1) (若第 x 維為 1，則所有 n<x 維都為 1)

所以最後 output layer 的 neuron, hyperplane 可定義為: $X_1 - X_2 + X_3 - X_4 + X_5 \dots + X_n = 0.5$ ① …… (第 x 項若 x 為奇則係數為 1，反之為 -1，假設 N 為偶)

第 X_m 代表上述 hidden layer 第 m 个 neuron 的 hyperplane 的 output

如上，我們已證明對 N 維 input 來說，N-N-1 network 絕對可以完成 XOR，因式 ① 的 hyperplane 只有在 input 有奇數个維度為 1 時值才會大於 0 (activate 之後就是 1)

(例: n = 6 維

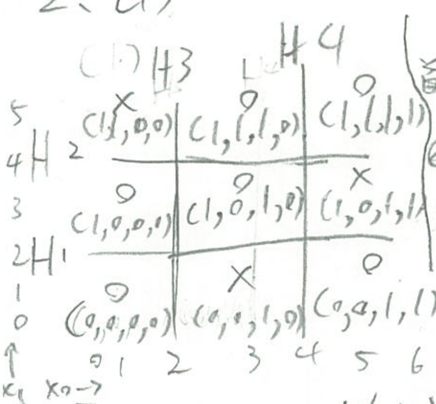
資料 (1, 0, 1, 1, 1, 0) 經 hidden layer ⇒ (1, 1, 1, 1, 0, 0)

再經 output layer, Activate(1 - 1 + 1 - 1 + 0 - 0 - 0.9) = 0

資料 (1, 0, 1, 0, 1, 0) 經 hidden layer ⇒ (1, 1, 1, 0, 0, 0)

再經 output layer, Activate(1 - 1 + 1 - 0 + 0 - 0 - 0.9) = 1)

2、(1)



首先把 first hidden layer 的 4 个 hyperplane 對應於 2 維 input 的 binary coding 寫出来。(以 hyperplane 的左方和下方為 negative，右方和上方為 positive。)，並且標上對應的 class label。
並且也列出可能的 2nd 3rd layer hyperplane 方程式来正確分類。
(H 代表 hyperplane，並且所有 neuron 都使用 hard limit 當 activation (大於 0 變 1，小於 0 變 0)

First layer: $H_1' \to x_1 = 2$, $H_2' \to x_1 = 4$, $H_3' \to x_0 = 2$, $H_4' \to x_0 = 4$

2nd layer: $H_1'' \to -0.5x_1 - 0.5x_2 + 0.5x_3 - 0.5x_4 = 0.3$

$H_2'' \to 0.5x_1 - 0.5x_2 + 0.5x_3 + 0.5x_4 = 1.3$

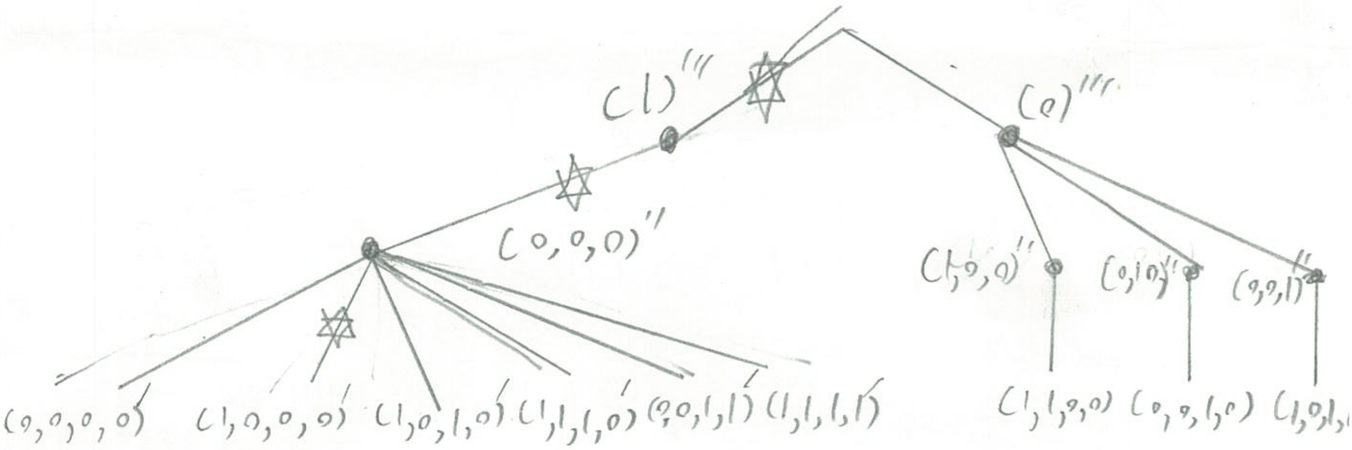
$H_3'' \to 0.5x_1 + 0.5x_2 - 0.5x_3 - 0.5x_4 = 0.8$

這三個 hyperplane 各自把三个對應 × label 的 pattern 獨立出来

3rd layer: $H''' \to -x_1 - x_2 - x_3 = -0.5$，$H'''$ 大於 0 對應 label ○，小於 0 對應 label X

(H''' 若 2nd layer 中不屬於任何 X pattern 就會是 1)

畫成 hidden tree 如下：



(2) (1,3) 對應 1st layer 的 (1,0,0,0)'，而若某一 pattern 為 ambiguous，其向上連結出去的 pattern 也都會是 ambiguous，thus (1,0,0,0)'，(0,0,0)''，(1)''' 都會混有 X label 的 data，如上圖 ✡ 号所示。