國立臺灣師範大學 106 學年度第 二 學期考試試卷 107年 5月 11日

得分：

1. (a) math and scentific computing, Graph, word processing

(b) learn, analyze, organize

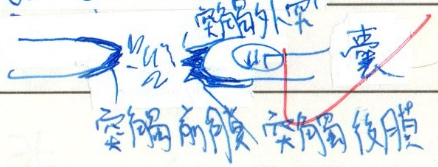
2. (a) brain, spinal cord, periphery -1

(b) organize, learn, comuting

(c) analyze, learn, feedback

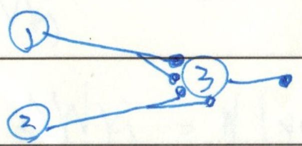
3. unipolar, bipolar, multipolar No address their function -1.5.

4. ① ~~encode~~ training ③ ~~training~~ production : recall the value -1.

5. 

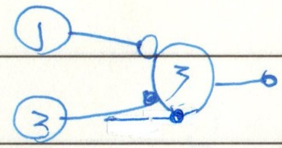
6. convergence, divergence, feedback

7. disjunction



$N3(t) = N_1(t-1) \vee N_2(t-1)$

conjoined negation



$N3(t) = \neg N_1(t-1) \wedge N_2(t-1)$

8. $N_3(t) = N_1(t-1) \vee a(t-1) = N_1(t-1) \vee (N_2(t-2) \vee b(t-2)) = N_1(t-1) \vee (\neg N_2(t-2) \vee N_2(t-3))$

$N4(t) = N_2(t-1) \wedge b(t-1) = N_2(t-1) \wedge N_2(t-2)$

9. 當A能激發B且能持續發射時，代謝~~射~~系統能激發1個或2個，這樣A的效率會提升 −5.

10. the system evolve over time

11. adaline: adatipative linear neuron

madaline: multiple adatipative linear neuron

12. ① Initialize the value $W(t_0)$

② Determine the steepest descent direction $-\nabla_w \xi(W(t)) = \frac{d\xi(W(t))}{dW} = 2(p^T W - R)$

③ find the weight

④ repeat 2~3

13. ∵XOR是比x_1, x_2，當x_1 False，x_2 True，代表不同，因此可以執行，然而 perceptron 的最後 circuit 有not 也沒有 not_{im}

因此無法完整用XOR表示 缺圖示 −3.

14. ① $\xi(W) = \langle \varepsilon_k^2 \rangle - W^T \langle X_k X_k^T \rangle W + \langle d_k X_k^T \rangle W = \xi - W^T R W + p^T W$ —(1)

② $\frac{d\xi(W)}{dW} = 2RW - 2p$ —(2)

$2RW - 2p = 0 \Rightarrow 2RW = 2p \quad W = R^{-1}p$ —(3) −3.

③ $d(k)^2 - p^T W$.

國立臺灣師範大學 學年度第 學期考試試卷 年 月 日

得分： 科目： 系級： 學號： 姓名：

15. ①要知道 R, P 的值，必須事先知道曲面

② steepest descent learning 是一個 batch method

16. advantage 較 steepest 來的容易 $f(w) = \frac{1}{2}aw^2 + bw + c$

disadvantage ~~只有 2 維，∴只有極好 or 極壞~~ −2.5.

17. minimum disturbance principle：最小擾動定律，~~當 training 時不能一起全部跑完~~ −4.

18. 因為每一層的輸出為下層的輸入，而 BP 是由最末層把值往第一層做更新

19. 一層一層的跑，不會有人先跑出結果 −3.

20. $E = \frac{1}{2}\sum_{k=1}^{L}(y_k - O_k)$

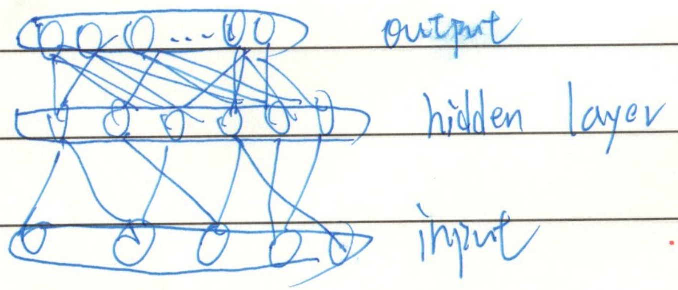
$$\begin{cases} -\frac{\partial E}{\partial W_{kj}} = \lambda[y_k - O_k]\,O_k(1-O_k)(x_j - V_{kj})^2 \\ -\frac{\partial E}{\partial V_{kj}} = \lambda[y_k - O_k]\,O_k(1-O_k)[2W_{kj}(x_j - V_{kj})] \end{cases}$$

$$\Rightarrow \begin{cases} \Delta W_{kj} = \eta'(y_k - O_k)\,O_k(1-O_k)(x_j - V_{kj})^2 \\ \Delta V_{kj} = \eta'(y_k - O_k)\,O_k(1-O_k)[-2W_{kj}(x_j - V_{kj})]' \end{cases}, \quad \eta' = \eta\lambda$$

-5

21. hidden layer 愈多層計算量大 , node是代表他們的feature或者他們之間的relationship

22.



23. (a) Mexican-hat function can be realized using Difference of Gaussians or Laplacian of Gaussian 圖?

8

(b) DOG (Difference of Gaussian) 算式?

LOG (Laplacian of Gaussian)

24. (i) LAM $\phi(b_i) = b_i$

definition & matrix memory representation?

(ii) CAM $\phi(b_i) = a_i$

-9

(iii) AM $\phi(b_i) = b_i + \varepsilon$

25. (a) $[b^1, b^2, b^3 \cdots b^M] \begin{bmatrix} 1 & 1 & 1 \cdots & 1 \\ 2 & 2 & \cdots & 2 \\ 3 & 3 & \cdots & 3 \\ \vdots & \vdots & \ddots & \vdots \\ M & M & \cdots & M \end{bmatrix} = [a^1, a^2, a^3, \cdots a^M]$

5

(b) $[b^1, b^2, b^3 \cdots b^M] \begin{bmatrix} 1 & 0 & \cdots & 1 \\ 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & 0 \\ 1 & \vdots & \cdots & 1 \end{bmatrix} = [a^1, a^2, a^3, \cdots a^M]$

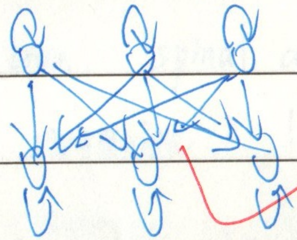
-24

國立臺灣師範大學 學年度第 學期考試試卷 年 月 日

得分： 科目： 系級： 學號： 姓名：

26. $\Sigma(2i-1)^2 = \Sigma(-(2i-1))^2$

-2. 代表 $2i-1$ 是 $-(-2i-1)$ 的 complement … 所以 …?

27.  ~~bidirection~~ associative memory

~~28.~~ Crosswalk phenomenon with a memory，從 A 到 B -2.

~~29.~~ 是一種 Dyamic system -5

30. $E = -Y^{T}WX = -WXY = -\sum_{i=1}^{m}\sum_{j=1}^{n}|W_{ij}|$ 解釋?

-3.