

(總分 100)

1. (5%) In 1950, Alan Turing proposed the concept of a universal machine (Turing machine), and proceeded the famous Turing Test. What is the Turing Test? What are the remark qualities that make it really universal?

2. (5%) AI 從 1956 年誕生，到 1970 年代的專家系統，為何有此發展？AI 發展做了何種改變？

3. (5%) 傳統 rule-based 專家系統問題有哪些？如何改善？

4. (10%) (1) 何謂 Bayesian rule (事後機率)？為何使用事後機率推論？舉例說明。

$$p(A|B) = \frac{p(B|A) \times p(A)}{p(B)}$$

(2) 使用 Bayesian rule 要注意那些條件(成立才可使用展開公式)？

$$p(H_i | E_1, E_2, \dots, E_n) = \frac{p(E_1, E_2, E_3, \dots, E_n | H_i) \times p(H_i)}{\sum_{k=1}^m p(E_1, E_2, \dots, E_n | H_k) \times p(H_k)}$$

5. (10%) 警察抓到四個嫌疑犯：甲、乙、丙、丁(涵蓋所有可能)，他們由前科紀錄猜測可能犯案機率分別為 0.3、0.2、0.1、0.4；警方又依序找到四項證物(彼此獨立無關)，分別是四位嫌犯留下之機率如下表所示：

留下之機率	甲 0.3	乙 0.2	丙 0.1	丁 0.4
證物 1	0.2	0.5	0.8	0.1
證物 2	0.5	0.4	0	0.4
證物 3	0	0.2	0.1	0
證物 4	0.4	0.1	0.2	0.5

請問：誰最有可能犯案？(以 Bayesian reasoning 求解，每一過程(當證物是陸續發現，每一證物出現後，誰最有可能，要說明)均要呈現，只寫答案不給分) 相關公式：

$$p(H_i | E_1, E_2, \dots, E_n) = \frac{p(E_1, E_2, E_3, \dots, E_n | H_i) \times p(H_i)}{\sum_{k=1}^m p(E_1, E_2, \dots, E_n | H_k) \times p(H_k)}$$

4.

(1) 事情發生後，探討原因的機率，可用來推論可能的原因。

(2) 使用專家系統時，利用事後機率推論可用來修正人類專家的推論(事前機率)，提高結果準確率。

9.

"輸入分佈的情況和其歸屬的程度值(輸出 Z)是否有符合題目說的成正比(是否符合初值假設)
(X, Y)

(2)

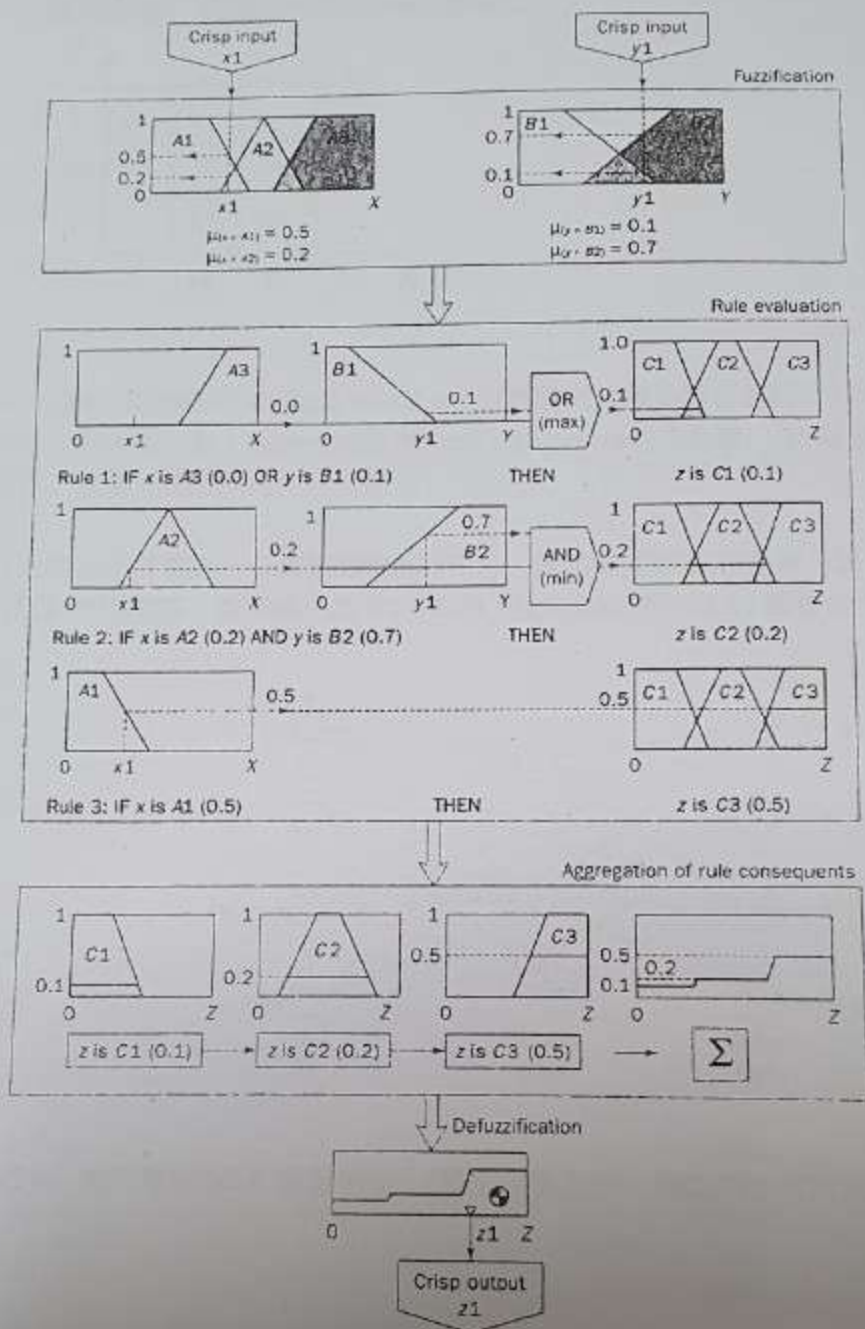
X 為 M 且 Y 為 S 那將不應為" RL "(沒有符合題目要求的正比)應該改為" RS "才有可能符合實際情況
= L = M = " M "

綜合以上 2 點可得出修正後的示意圖 =

" RL "

6. (5%) 模糊理論中，何謂(1)語意變數(linguistic variables)及對應之語意值？
(2)fuzzy set 如何表示？以天氣溫度為例，作一說明。

7. (10%) 下圖為 Mamdani 模糊推論方式，請設計一範例，有四條規則，每一規則則有兩條件 $X \cdot Y$ ，輸出為 Z ，分別有兩規則用 AND 連結，兩規則用 OR 連結，說明如何運作。(用圖示與公式說明)



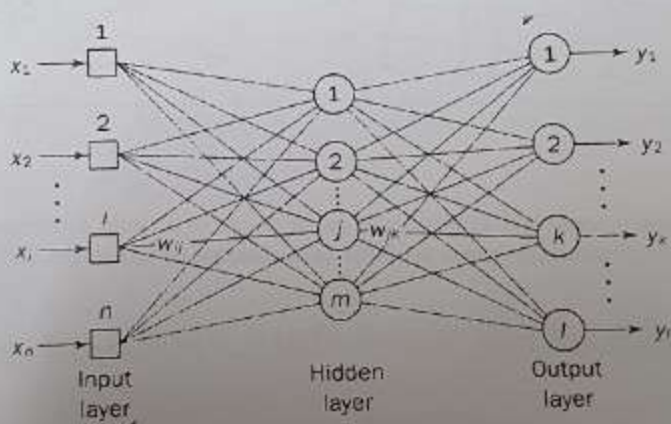
8. (5%) 發展一套 fuzzy expert system，最重要步驟為何？哪一步驟將會花費最多時間？
 解問題為何，訂定適合的規則 調整 (set fuzzy set)

9. (10%) 下列模糊規則建立原則為：輸出 Z 與輸入 X, Y 均成正比，輸入模糊集合(大：L，中：M，小：S)，輸出模糊集合(大：L，有些大：RL，中：M，有些小：RS，小：S)，其建立的模糊關聯記憶體(FAM)如下，如何看出此 FAM 是否合理？你認為何處不合理？請更正。

Y				
L	M	RL	L	
M	RS	M	M	
S	S	RL	M	
	S	M	L	X

10. (10%) 類神經網(artificial neural net)以何種方式達到學習目的？以訓練方式區分，可大致分為：supervise(教導) and unsupervised(不教導)，此兩種方式如何運作？並舉一代表模式。
 監督式學習：教師提供目標輸出，模型學習如何產生該輸出。
 非監督式學習：模型學習數據的內在結構，無需目標輸出。

11. (10 %) 下圖為一多層認知模式(MLP)，請問如何學習(假設有 10 筆樣本(資料)X1~X10 要學，對應輸出為 Y1~Y10)，說明如何學習這 10 筆樣本？



12. (5%) 如何增快 MLP 學習速率？(舉出三種方法，並說明如何計算)

13. (10%) 下圖為 2 inputs & 100 output neurons ($= 10 \times 10$) SOFM 訓練過程。(a) 初始權重, (b) 訓練 100 次, (c) 1000 次, (d) 10000 次結果, 請問:

1. 圖中黑點意義為何? 線條意義為何?
2. 為何圖會如此變化? 原因為何?

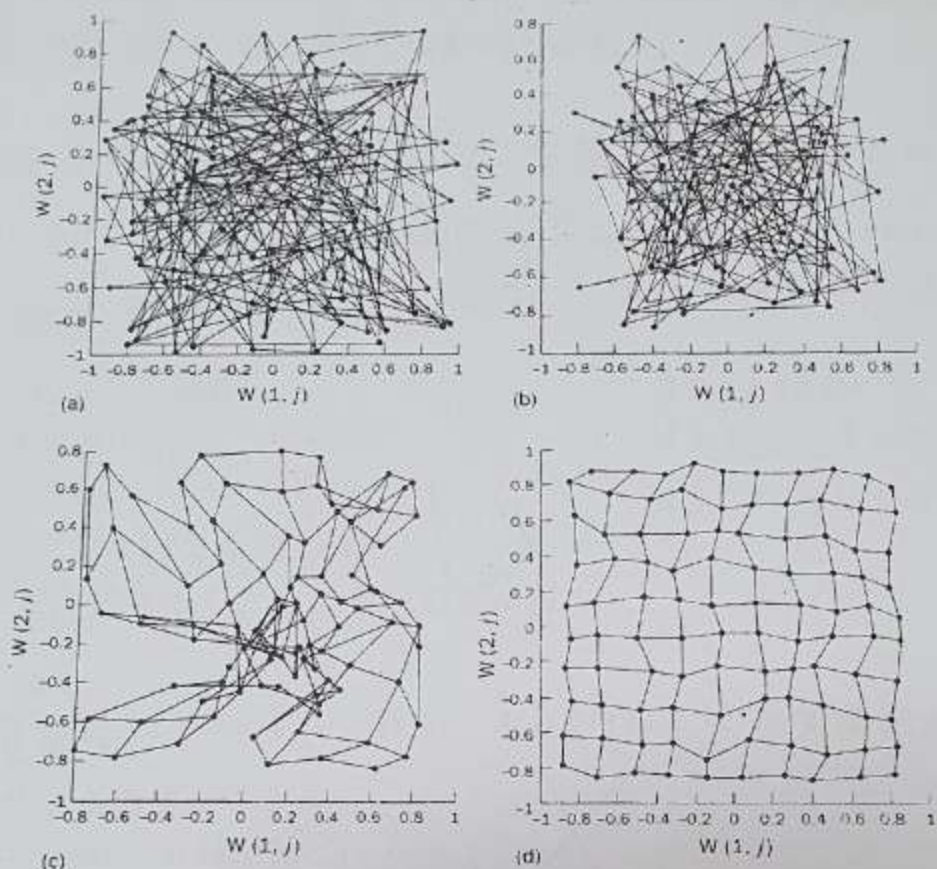


Figure 6.28 Competitive learning in the Kohonen network: (a) initial random weights; (b) network after 100 iterations; (c) network after 1000 iterations; (d) network after 10,000 iterations

- (1) 黑點為權重的分佈情況: 線條兩端的點代表著在 neural network 中兩點位置相鄰
- (2) 機器在訓練的時候沒被給予標準的答案, 只提供範例資料, 機器會自動找尋潛在的規則, 權重的分佈會趨於常態 (由大到小, 由左到右, 每筆資料都有其對應點), 會愈接近真實的分佈情形。

3.

(1) 事件發生後，計算每一種可能發生原因的機率；所以，公式是說明觀察到事件 E_1, E_2, \dots, E_n 發生，

看是 H_i 造成之機率，可由事前機率 $P(E_1, E_2, E_3, \dots, E_n | H_i)$ 與 $P(H_i)$ 去推算事後機率 $P(H_i | E_1, E_2, \dots, E_n)$ 。
公式中， H_i 必須互斥 (mutual exclusive) 且完整涵蓋所有可能， E_i 必需條件獨立分子才能展開。

$$P(E_1 E_2 E_3 \dots E_n | H_i) = P(E_1 | H_i) P(E_2 | H_i) P(E_3 | H_i) \dots P(E_n | H_i)$$

(2) 不行使用貝氏。因為，疾病觀察到的症狀 E_i 無法條件獨立，例如，感冒造成 = 咳嗽、喉嚨痛與流鼻涕，都有關聯，無法條件獨立。另外，可能疾病 H_i 無法互斥 (mutual exclusive) 且不可能涵蓋所有可能，例如 = 造成咳嗽可能疾病，有流感、肺炎、支氣管炎、感冒等。彼此沒有互斥，而且，無法保證列出所有可能疾病。

① 修改/調整權重以適應
② supervise = 針對時期

unsupervised



國立臺南大學

學年度 考

考試科目:

AI

班 別

學

主 名

案卷

總分

61

①

Turing Test 分為 2 個階段。

第一階段為將一男一女關在房間透過電腦與房外溝通，男生要假裝自己是女生，女生要證明自己是女生，看看房外的人是否可分辨出 2 人性別。

第二階段為將男生用電腦替代，看看是否能分辨出兩者。

② 詢問者可詢問跨領域的問題，具有客觀的標準。 *測試與系統無關

2

1970 年代初期，研究 AI 人員專注於開發解決問題的一般方法，早期的程式應用搜尋策略測試小步驟的組合，再試圖擴大解決大問題，但這樣是錯的，而且許多 AI 欲解決的問題太龐大難，美英政府支持的專案皆因沒有明顯的成效而被迫中斷。

1970 年代後期，AI 研究者逐漸了解唯有解決一些典型案例或擁有一些成效才能受到重視，因此有了 DENDRAL 專家系統，DENDRAL 的出現也代表著 AI 的一些改變，但專家系統有以下缺點：① 專家知識的定義過於狹隘而受到限制 ② 對使用者的廣泛需求較不健全且不彈性 ③ 解釋能力有限 ④ 難被驗證，現今沒有通用的技術分析其一致性和完整性 ⑤ 尤其第一代無法從自己的經驗中校正，故無法發展快速，所以轉而發展機器學習。

(1) 在使用專家系統時，利用事後機率的推論來修正人類專家的推論（事前機率），提高準確度。 $H_1 \sim H_n$ 必須互斥且完整涵蓋； $E_1 \sim E_n$ 必須條件獨立。

(2) 不行。因為我們無法確定觀察到的症狀 (Evidence) 間是否彼此獨立，即使不同症狀也有可能由相同的致病因子。

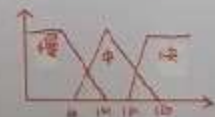
(1) linguistic variables

linguistic value

hypothesis

速度

快 (快中快)



Step 1 = fuzzification 模糊化：將 crisp input 轉成模糊集合，賦予適當歸屬程度。

Step 2 = Rule evaluation 規則評估：透過 OR 或 AND 操作，獲得輸出權重。

Step 3 = Aggregation of rule consequents 重疊：將 Step 2 的輸出結果合併成一個。

Step 4 = defuzzification 解模糊化：將 Step 3 的結果利用重心法，轉成一個 crisp output (重心法：半圓形之重心)。

5.

① 修改/調整權重以達學習目的。

② supervise = 針對每一輸入事件先定好預期輸出樣本，再由 input 和 output 與期望輸出做平方誤差，進行權重修改，讓實際輸入和預期輸出的誤差縮小。如：MLP

unsupervised = 權重之修改完全由 input 與當下每一 output 之權重比較，求出一 winner 再以 winner 及其四周做權重修改，事先沒有做任何指派 output 或計算權重之動作。如：SOFM

6. MLP =

Step 1 = Initialization 初始化 = 利用均勻分佈在小範圍的亂數來設定初始值和臨界值。

Step 2 = Activation 活化 = 用實際輸入 x_i 和期望輸出 y_i 計算

① 隱層神經元的實際輸出 z_i

② 輸出層神經元的實際輸出 y_k

Step 3 = Weight training 權重訓練 = 計算隱層、輸出層的誤差梯度 (如 δ_i)

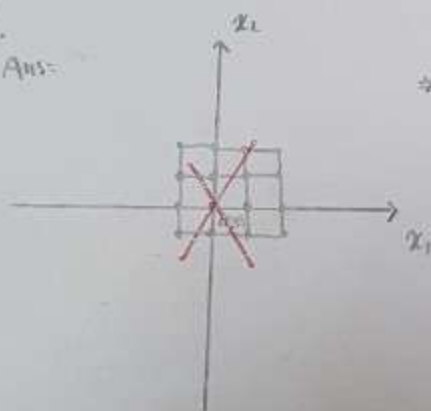
來調整權重。 * 利用偏微(誤差)來調整權重

Step 4 = Iteration 迭代 = 迭代次數 PI ，回到 step 2 重複過程直到符合中止條件為止。

— 1 —

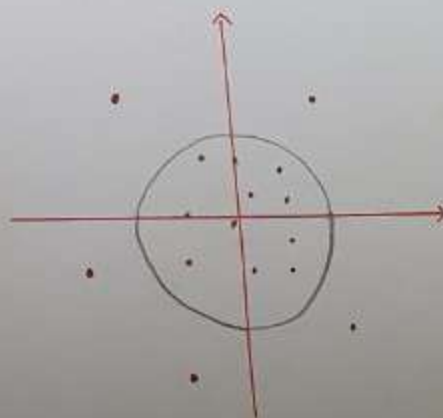
7.

Ans:



* input 的特徵(資料分佈)會呈現在 output neuron 的 weight 上。

— 10 —



8.

certainty factor = $\frac{\max\{0, 1 - \min\{0, 1 - \text{cf}_i\}\}}{1}$ 1° 當 $i=1$ 時, $\text{cf}_1 < 1$ 且 > 0 成立 $\rightarrow \text{OK}$ 2° 假設 $i=n-1$ 時, $\text{cf}(cf_1, cf_2, \dots, cf_{n-1}) < 1$ 成立3° 因為 $\text{cf}(cf_1, cf_2, \dots, cf_n) = \text{cf}(cf_1, cf_2, \dots, cf_{n-1}, cf_n) = \text{cf}(cf_1, cf_2, \dots, cf_n) + \frac{cf_n \times (1 - cf_1) \times (1 - cf_2) \times \dots \times (1 - cf_{n-1})}{1 - cf_1 - cf_2 - \dots - cf_{n-1}}$ 所以當 $i=n-1$ 的假設成立時可看成 $\text{cf}(x, y)$, $\begin{cases} x = \text{cf}(cf_1, cf_2, \dots, cf_{n-1}) < 1 \\ y = cf_n < 1 \end{cases}$ x 與 y 皆小於 1 大於 0, 因此由數學歸納法可得出: $\text{cf}(cf_1, cf_2, cf_3, \dots, cf_n) < 1$ 得證 $x = (1 - cf_{n-1})$
中每一規則 cf_i 皆 > 0
 $< 1 \therefore \text{cf}_n$ 只需愈乘
愈小 $> 0 < cf_n < 1$

$$x < 1, y < 1 \rightarrow x + y < 1 ?$$

-10

Ans = 若 n 個規則, 其最大 $\text{cf} = p$, 則

$$\text{cf}(cf_1, cf_2, cf_3, \dots, cf_n) = cf_1 + cf_2 \times (1 - cf_1) + cf_3 \times (1 - cf_1) \times (1 - cf_2) + \dots + cf_n \times (1 - cf_1) \times (1 - cf_2) \times \dots \times (1 - cf_{n-1})$$

$$< p + p(1-p) + p(1-p)(1-p) + \dots + p(1-p)(1-p) \dots (1-p) \text{ (共 } n \text{ 個)} = (p(1-(1-p)^n)) / (1-(1-p)) < (p(1-(1-p)^n)) / (1-(1-p)) = p/p$$

 $= 1$ 所有 n 個規則的 certainty factors 總和 < 1 .