

行政院國家科學委員會補助專題研究計畫成果報告

動態貝式網路的建構與學習機制 － 航太業績效指標管理之應用

計畫類別： 個別型計畫 整合型計畫

計畫編號：NSC 90 - 2212 - E - 029 - 001 -

執行期間： 90 年 8 月 1 日 至 91 年 7 月 31 日

計畫主持人：王偉華

共同主持人：

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

執行單位：東海大學工業工程與經營資訊學系

中 華 民 國 91 年 10 月 31 日

行政院國家科學委員會專題研究計畫成果報告

國科會專題研究計畫成果報告撰寫格式說明

Preparation of NSC Project Reports

計畫編號：NSC 90-2211-E-029-001

執行期限：90年8月1日至91年7月31日

主持人：王偉華 東海大學工業工程與經營資訊學系

計畫參與人員：

協同主持人：廖仁傑 建國技術學院工業工程與管理系

研究生：蘇俊和 東海大學工業工程與經營資訊研究所

摘要

企業在面對全球化的競爭，如何維持長久的競爭力是倍受關切的，尤其當競爭者技術達到一定水準時，獲利來源難以再依靠先進的技術與製程，而是端賴有效率的管理企業的營運，以及預防管理機制的建立。在一些不良徵狀出現時，即能預知可能產生的問題；或發生問題時，可以迅速推測出可能影響的範圍與診斷問題產生的原因，並提前做好防範措施以減少產生損失的機會。

因此，對企業而言，建立一個可以總覽企業營運的績效評估系統，且具有回饋的學習機制，一直是企業管理者重視的焦點。一般而言，企業內部之各部門都會設定其自有之績效指標並竭盡所擁有的資源去達成。然而各部設定之績效指標間並非存在全然獨立的關係，經常產生相互牽制之效應。因此，為了有效整合企業內林林總總之績效指標，使各部門努力目標與組織目標趨於一致，必須將企業內部之績效指標依其因果關係結合成為一個績效指標網(Performance Index Network; PIN)。

在本研究中，結合領域專家知識與貝氏網路之相關技術為基礎，發展一個可以處理企業動態環境之不確定性及時間延遲的知識管理工具 - Temporal Node Bayesian Network。並將此知識管理工具運用在航太製造業，建立一個可以達到預測與診斷功能的績效指標系統，作為本研究之實務驗證。

關鍵詞：績效指標網，貝氏網路，Temporal Node Bayesian Network，預測與診斷

Abstract

Upon to the intensive competition, enterprise should not only to gain the competitive advantages but also sustain the competition. However, due to the fast advanced technology movement, the competence of an enterprise is partly shifted to the capability of aligning her operations management with the strategies and the mechanism of diagnosis and prediction the operational processes. The capability of knowing the possible catastrophes and knowing the causes of those catastrophes becomes the winning features of any success enterprise.

Performance Indexes (PIs) Management has been adopted along with this paradigm. However, those PIs were used to be managed independently. Most of the enterprises were suffered by the conflicting PIs. In this research, we believed that all the PIs should be considered, somewhat, dependently for goal congruence, and thus, proposed a Performance Index Network (PIN) architecture.

In this research, we combined the domain expert knowledge and the Bayesian network methodologies to develop the temporal and dynamic mechanism. A real aerospace industry application was adopted to justify the result of the proposed method.

Keywords: Performance Index Network, Bayesian Network, Temporal Node Bayesian Network, diagnosis and prediction.

一、研究動機與目的

企業在面對來自全球遽烈的競爭環境下，任何一個計畫管制作為都必須輔以良好之監、控機制以及相關單位的配合來完成，而監控機制之發揮則反映在即時(Real-Time)與連線(On-Line)之績效指標系統上。一般而言，企業內部之各功能組織都會設定其自有之績效指標並竭盡所擁有資源去達成它(們)。然而，各部門所設定之績效指標間並非存在全然獨立之關係，在企業有限之總體資源下，各部門所設定之績效指標更可能產生互相牽制之效應。例如飛機製造業中，零件製造單位以生產數量之提昇為績效衡量方式，而裝配線上則以準時交機為努力目標，在各自之績效目標驅動下，往往製造單位員工因追求生產數量之成長而挑選容易加工之零件製作，而在裝配線上卻遲遲無法等到完成裝配整架飛機所需之關鍵性(難度高)的零件，其結果當然導致飛機之交期延誤及公司之連帶損失。因此，為了有效整合企業內林林總總之績效指標，使各個部門之努力目標與組織目標趨於一致(Goal Congruence)，必須將企業內部所有之績效指標依其因果關係與互為影響的程度釐清，結合為一個具因果關係的網路圖形，成為一個適當地反映現況的績效指標網(Performance Index Network ; PIN)。

建構績效指標網之方法有許多，其一為透過領域專業人士(Domain Expert)之主觀認知直接將績效指標網勾勒出來；另一種則為透過對取得資料之客觀分析而來（即利用資料探勘之技術）。在本研究中，我們同時利用機器學習(Machine Learning)和貝氏網路(Bayesian Network)之相關技術為基礎，發展一個可以同時處理企業動態環境之不確定性及時間變異性之知識管理(Knowledge Management)工具 – Temporal Node Bayesian Network。

本研究以航太業為例，利用 Temporal Node Bayesian Network 所具有之時間序列 Data Mining 功能及學習機制，建立一個可達到監測、控制、配合功能之績效指標網路。

二、研究方法

建構定態性貝氏網路的方法有兩種，其一為透過領域專業人士的主觀認知將貝氏網路勾勒出來，另一種則是透過取得客觀的資料分析而來。然而對於具時間延遲特性的問題，定態性貝氏網路似乎無法貼切的表達變數間在時間上的因果關係，以及難以精準推論變數間相互影響的程度。在本研究中建構擷取動態機制的貝氏網路的方式為結合上述兩種方式：先透過資料分析，將資料所隱含的架構自動產生貝氏網路，再由領域專家修正其架構，使其較為符合問題領域的架構如圖 1.1 所示。

接著再透過設計擷取動態特質機制的定態性貝氏網路，替代動態性貝氏網路如圖 1.2，之後隨著資料量的增加，貝氏網路透過學習的機制，修正其架構與參數，使系統可以適切地描述問題領域。

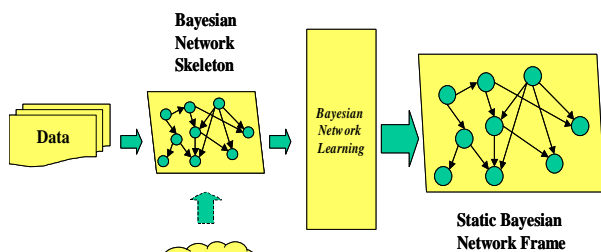


圖 1.1 貝氏網路的建構方式示意圖

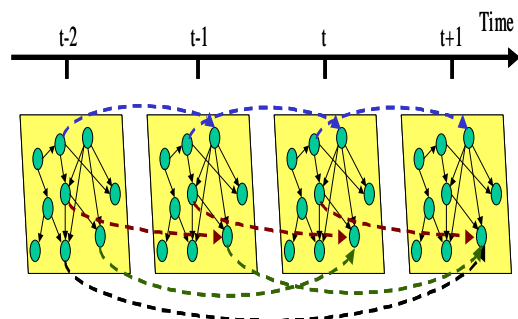


圖 1.2 動態性貝氏網路示意圖

三、系統設計

本研究提出結合兩種建構定態性貝氏網路的方式來建立初始的貝氏網路架構。也就是先利用資料

分析將資料所隱含的架構利用資料探勘的方式自動建構出來，再利用問題領域專家的專業來修正自動建構的貝氏網路，使其架構一開始就能有較好的表現，之後以其學習模式來修正貝氏網路的參數與架構，最後能得到一個較為符合問題領域現況的穩定狀態，如圖 3.1 所示。

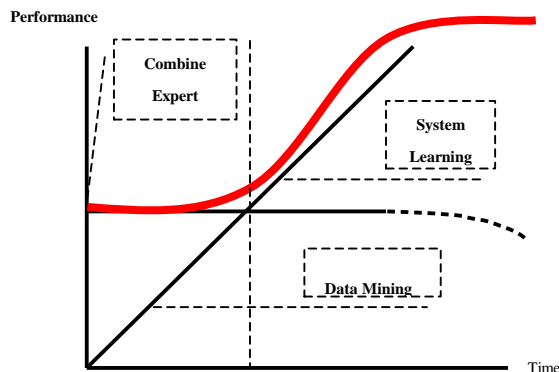


圖 3.3 定態性貝氏網路建構方式預期成效圖

本研究所初步設計定態性貝氏網路運用在航太產業績效評量系統之系統架構圖如圖 3.2 所示：

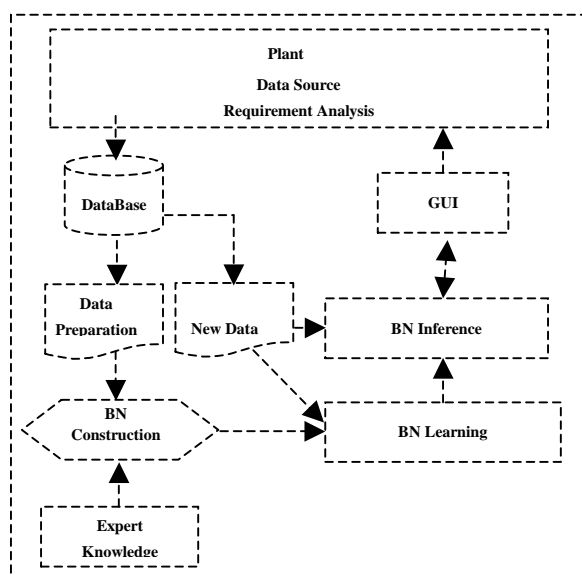


圖 3.4 定態性貝氏網路運用在航太產業績效評量系統架構圖

貝氏網路的學習分為架構上的學習(Structure Learning)與參數上的學習(Parameter Learning)。架構上的學習意指在找出最符合資料所隱含的貝氏網路結構。利用資料探勘的方式所找出的貝氏網路結構可能有很多個，此時搭配評量網路品質優劣的函數(Scoring Function)，給予分數的評量，用以找出最符合資料的結構，在本研究中所使用的評量函數為 BIC 法，其數學式如下：

$$BIC(B : D) = L(B : D) - \frac{\log N}{2} \#(G) \quad (3.1)$$

其中 $L(D : B) = Pr(D | G, \hat{\theta}_G)$

N 為 Cases 的數量；#(G) 為變數節點的數目
參數上的學習又稱為模式參數的擬合(Model

Fitting), 意指在貝氏網路結構已知下, 利用資料進行變數間聯合機率分配的計算, 使可以找出在這些資料中所代表的資訊, 分析出這些變數在此資料中的因果關係影響程度的行為是如何。隨著資料的增加與更新, 則進行變數間聯合機率分配的修正與學習, 使得與這些資料隱含的分配相符。本研究採用 EM Algorithm 做為參數上的學習的方法。

四、定態性貝氏網路系統建置與實作

本研究定態性貝氏網路績效指標系統的建置, 主要以 Hugin 公司開發的軟體 Hugin Researcher 為系統主體, 建構漢翔公司機工處中小型零件加工中心績效指標系統。

本研究期望透過所提出的機制, 建構出一個具診斷與預測功能的績效評估系統的原型, 具體改善目前績效評估方法。實驗將分二個部份進行:

1. 驗證結合資料探勘與專家知識兩種方式所建構出的貝氏網路其結構在初期時優於資料探勘所建構出的貝氏網路。

2. 驗證學習過程對於貝氏網路是重要的。

在績效衡量公式方面, 本研究期望工作中心的管理者在使用本研究設計的績效評估系統監測與控制整個工作中心的運作時, 能產生正確的預測與診斷的資訊, 提供管理者做為管理上參考的依據, 因此根據一些證據做出正確的推論為本研究衡量的指標, 所以本研究定義正確率為實驗驗證衡量指標, 其公式如下:

$$Performance\ Measure = \frac{X}{X+Y+Z} \quad (4.1)$$

X: 推論正確的資料筆數

Y: 推論不確定的資料筆數

Z: 推論錯誤的資料筆數

若 Performance Measure 的值愈大, 代表推論的正確率愈高, 則其愈能反映工廠生產的真實績效。

本實驗蒐集了漢翔公司機工處中小型零件加工中心 15 部加工機台於 90 年 10、11 月的生產資料, 包含了投入資源、各項指標資料, 以一天三班為例, 共蒐集到 131 筆資料。

將選取加工機台的資料以 50%、25%、25% 依順序分為三群, 以前兩群作為訓練資料(Training Data), 全部資料為作測試資料(Testing Data)。

【實驗一】: 驗證結合資料探勘與專家知識兩種方式所建構出的貝氏網路其結構在初期時優於只以資料探勘建構出的貝氏網路。

步驟一: 以資料探勘的方式自動建構貝氏網路

將加工機台的資料, 以第一群 50% 的資料作為自動建構貝氏網路的資料。

步驟二: 測試該貝氏網路的推論能力

當貝氏網路的結構建立後, 接著進行該網路結構的參數學習。以相同第一群 50% 的資料以 EM Algorithm 進行該網路結構分配的擬合, 求得第一群 50% 資料的聯合機率分配, 再以全部的資料進行該網路推論能力的測試。

步驟三: 結合專家知識的貝氏網路

由資料探勘的方式所建構的貝氏網路為與用來建構的資料相符合的結構, 並非真實的結構, 隨著資料的增加, 貝氏網路的結構與參數經過學習與修正的過程, 為了縮短此一過程, 適當地加入專家知識則有助於績效的改善。因此藉由此工作中心管理人員的領域知識, 修正了此一網路結構。

步驟四: 利用修正後的結構再次測試推論的正確性: 重覆步驟二的測試過程, 利用全部的資料數測試該結構推論的正確性評估。

步驟五: 假設檢定

檢定一: 檢定以前 50% 的資料做為訓練資料, 全部資料做為測試資料時, 資料探勘所建立的模型推論正確率與結合專家知識的模型推論正確率是否有差異。

【實驗二】: 讓初始資料面建構的貝氏網路, 隨著資料的增加與更新經歷學習與修正結構、參數的過程, 改善推論的正確率。

步驟一: 以第一群 50% 的資料加上第二群 25% 的資料自動建構貝氏網路。以前 75% 的資料作為建構貝氏網路的資料, 比只以第一群 50% 的資料建構的貝氏網路圖多了缺勤率影響機器稼動率以及下料工時會影響稼動率的因果關係。

步驟二: 測試貝氏網路的推論能力。當結構建立完成後, 接著進行前 75% 資料參數的學習, 之後再以全部的資料進行貝氏網路推論正確率的過程。

步驟三: 利用結合專家知識所建立的貝氏網路進行推論能力的評估: 利用結合專家知識所建立的貝氏網路結構, 經歷第一群、第二群 75% 的資料分配的擬合後, 接著利用全部的資料進行推論正確率的評估。

步驟四: 假設檢定

檢定二: 檢定以前 75% 的資料做為訓練資料, 全部資料做為測試資料時, 資料探勘所建立的模型推論正確率與結合專家知識的模型推論正確率是否有差異。

檢定三: 利用資料探勘自動建構的模型, 檢定以前 50% 的資料做為訓練資料, 全部資料做為測試資料與以前 75% 的資料做為訓練資料, 全部資料做為測試資料, 其推論正確率是否有差異。

由實驗一與實驗二之結果得知, 結合資料探勘與專家知識所建構的貝氏網路, 可以在系統初始時, 就能有較好的表現, 縮短了系統學習與修正的過程; 若資料量增多, 涵蓋問題領域的部份愈廣, 則用來建構出的網路結構愈能代表問題問題領域, 其推論正確率愈高。

然而, 實驗結果雖可說明結合專家知識的系統可以改善系統推論的表現, 但整體而言其結果並非很好, 所以以定態性貝氏網路結構來解決具時間延遲的績效評估問題, 似乎不是很恰當。因此針對具時間延遲與具不確定性的問題, 本研究進一步提出具擷取動態機制的貝氏網路來解決。

五、擷取動態特性貝氏網路的建構

由於工廠在生產作業時的生產狀況具時間延遲性與作業員報工不確實，造成績效評估的不易，投入的生產資源其產出並非在同一個時間區段中立即可見。往往延遲產出在不同的時間區段，因此其投入與產出的因果影響性並非只在一個時間區段內，而是跨越了不同的時間區段，針對時間延遲的問題。一般以定態性貝氏網路的模式，並不能有效的解決，因此對於具時間延遲且不確定性的問題多數是以動態性貝氏網路模式加以善，用以彌補定態性貝氏網路之不足。然而動態性貝氏網路隨著變數數量的倍增，使得網路複雜度成爆炸性的成長，造成運作系統的作業平台運作資源不足，以及運作時間的耗費。本研究則提出以具時間關連變數的定態性貝氏網路 (Temporal Node Bayesian Network) [1]，替代動態性貝氏網路，用以解決具時間延遲與不確定性的問題。

為了有效簡化動態性貝氏網路所形成的狀態空間以及降低網路的複雜度，本研究提出以具時間關連性變數的定態性貝氏網路結構 (如圖 5.1) 替代動態性貝氏網路，用以解決具時間延遲與不確定性的問題。

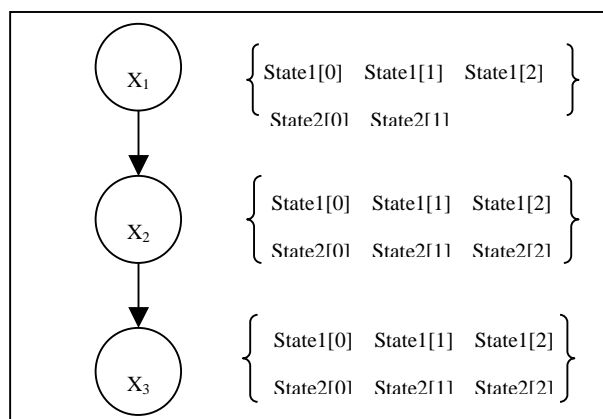


圖 5.1 具時間關連性變數定態性貝氏網路架構圖

本研究針對漢翔公司機工處中小型零件加工中心的績效評估系統進行檢討，發現影響因子變數所造成的影響，並非在同一個時間區段，即由績效指標變數顯示出，其所造成的影響，往往會有部份影響落後延遲的現象發生。因此本研究針對此落後延遲的現象，設計一處理動態機制的貝氏網路架構圖，如圖 5.2 所示：

為進行具時間關連性變數的定態性貝氏網路系統實作，本網路結構的建立主要是以定態性貝氏網路為基礎，其變數間的因果關係如同定態性貝氏網路，在整個網路結構中視每個班別的影響因子變數群為獨立變數群，並將班別內的影响因子變數做分群的處理。當具時間關連性變數的定態貝氏網路建立完成後，則進行參數的學習，如同定態性貝氏網路的學習模式，在本研究中仍是以 EM Algorithm 做為求得聯合機率分配的演算法。

【實驗三】：以具時間關連性變數的定態性貝

氏網路測試本研究績效指標系統的推論正確性評估。

步驟一：對具時間關連性變數的定態性貝氏網路進行參數的學習。如同實驗二，先將資料以 75%、25% 分成兩群，前 75% 的資料做為參數學習的訓練資料，再以全部的資料用 EM Algorithm 進行參數的學習。

步驟二：測試具時間關連性貝氏網路的推論能力

步驟三：假設檢定

檢定四：以二個獨立母體比例之右邊檢定，檢定具時間關連性變數的定態性貝氏網路的推論正確率比定態性貝氏網路的推論正確率高：

由實驗三的檢定結果可知，對於處理時間延遲性與不確定性的績效評量問題，具時間關連性的定態性貝氏網路的推論正確率與定態性貝氏網路的推論正確率有顯著差異，亦即具時間關連性變數的定態性貝氏網路其推論正確率要比定態性貝氏網路推論正確率要高。

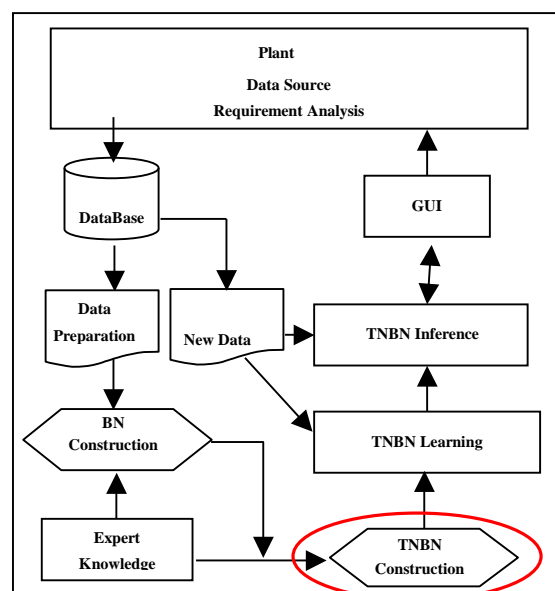


圖 5.2 具時間關連性變數的定態性貝氏網路運用在航太產業績效評量系統架構圖

六、結論與未來研究方向

根據本研究的實驗結果，得到下列結論：

1. 透過資料探勘與專家知識的結合確實能讓所建立起的績效指標網於短時間內提高系統的效能。
2. 對於具時間延遲性的問題，若以動態性貝氏網路來描述問題，則會造成動態性貝氏網路結構過於複雜，模型大小隨著變數的增多與狀態空間的擴大成爆炸性的成長，大大地降低了系統運作的效能，針對此問題，本研究提出以具時間關連性變數的定態性貝氏網路來處理，一來有效降低動態性貝氏網路的模型複雜度，二來改善了定態性貝氏網路處理具時間延遲性與不確定性的問題的表現。
3. 利用貝氏網路的技術解決工業界的真實問題

時，由於問題複雜度的因素，會使得系統遭遇到記憶體不足以及運作效能偏低的問題。尤其是問題領域變數的狀態時，對記憶體空間的需求會成爆炸性的成長。

在建議與未來研究方向方面，由於本研究提供一個結合資料探勘與專家知識的機制來處理具時間延遲性與不確定性的問題，雖然其在效果上改善了定態性貝氏網路的效果，但其效果並非很顯著，除了對學習模式加以探討外，仍有一些問題值得進一步的探討，使得本研究趨於完善。本研究提供以下建議供後續研究者參考：

1. 資料處理：由於本研究是以資料分佈的範圍等距設定為幾個狀態，這種狀態設定的方式是否恰當，值得商榷。因此狀態設定若能依目標導向(Goal Oriented)設定，也就是資料數值分布較少的範圍，其狀態設定愈放寬，資料數值分布愈多的範圍，則其狀態設定愈細分，如此推論效果更將顯著。
2. 變數的分群：在本研究驗證過程中，發現若不將變數分群，則其所造成的貝氏網路結構的狀態空間會呈現爆炸性的成長，若能將變數分群，則能有效降低其模型狀態空間的大小，同時增進模型的效能。因此變數如何分群是一個值得研究的方向。
3. 動態機制擷取的回饋機制：本研究中提出以具時間關連性變數的定態性貝氏網路來取代動態性貝氏網路結構過於複雜的現象，但其並無回饋修正的機制，因此如何建立回饋修正的機制，是一個相當值得研究的重點。

五、參考文獻

- [1].Arroyo-Figueroa, G., Sugar, L.E., and Villavicencio,A., “Probabilistic temporal reasoning and its application to fossil power plant operation.” *Expert System with Applications* 15(1998) 317-324.
- [2].Barrientos M.A., Vargas J.E., “A framework for analysis of dynamic processes based on Bayesian networks and case-based reasoning.” *Expert System with Applications* 15(1998) 287-294.
- [3].Bilmes, Jeff(1998). A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov model. *International Computer Science Institute*.
- [4].Enrique, C., Jose, M.G., and Ali S.H., (1997)*Expert Systems and Probabilistic Network Models*, Springer-Verlag, New York.
- [5].Friendman, N., K. Murphy, and S. Russell. (1998) Learning the structure of dynamic probabilistic networks. In *UAI*.
- [6].Friendman, N. (1997). Learning belief networks in the presence of missing values and hidden variables. In *ICML-97*.
- [7].Ghahramani, Z. and S.T. Roweis. (2001) Learning nonlinear dynamical systems using an EM algorithm, <http://www.gatsby.ucl.uk>.
- [8].Heckerman, D. A tutorial on learning with Bayesian networks. *Technical Report MSR-TR-95-06*, Microsoft Research, March, 1995 (revised November, 1996).
- [9].Heckerman, D. and D. Geiger. Learning Bayesian Networks: The Combination of Knowledge and Statistical Data. *Technical Report MSR-TR-94-09*, Microsoft Research, March, 1994 (revised February, 1995).
- [10].Jaakkola, F.(1996). “A introduction to Bayesian network.” Springer.
- [11].Jordan, M. and R. Jacobs (1994). Hierarchical mixture of experts and the EM algorithm, *Neural Computation*, 6, p.181-214.
- [12].Kjaerulff, Uffe(1995). dHugin: a computational system for dynamic time-sliced Bayesian network. *International Journal of Forecasting*, 11: p.89-111.
- [13].Lauritzen, S., Thiesson, B., and Spiegelhalter, D.(1994), Diagnostic systems created by model selection methods: A case study, In Cheeseman, P. and Oldford, R., editors, *AI and Statistics IV*, volume *Lecture Notes in Statistics*, 89, pages 143-152. Springer Verlag. New York.
- [14].Long, William, “Temporal Reasoning for Diagnosis in a Causal Probabilistic Knowledge Base” Reprinted from *Artificial Intelligence in Medicine*, 8: 193-215, 1996.
- [15].Nicholson, A.F., Brady J.M., “Dynamic Belief Networks for Discrete Monitoring.” December 20, 1993.
- [16].Nikovski, Daniel “Learning Stationary Temporal Probabilistic Network.” The Robotics Institute, Carnegie Mellon University.
- [17].Pearl, J. (1986). Fusion, propagation, and structuring in belief networks. *Artificial Intelligence*, 29: 241-288.
- [18].Pearl, J. *Causality: Models, Reasoning and Inference*. Cambridge University Press.
- [19].Schervish, M(1995). Probabilistic inference and Inference diagrams. *Operations Research*, 36: 589-604.
- [20].Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* 6, 461-464.
- [21].Shachter, R. (1988). Probabilistic inference and inference diagram. *Operations Research*, 36: 589-604.
- [22].Singh, Moninder and Valtorta, Marco “Construction of Bayesian Network Structures From Data: A Brief Survey and an Efficient Algorithm.” *International Journal of Approximate Reasoning* 1995; 12: 111-131.

- [23].Spirtes, P., C. Glymour, and R. Scheines (1993).
Causation, Prediction and Search. No. 81 in lecture
Notes in Statistics. New York: Springer-Verlag.
- [24].<http://www.hugin.com>
- [25].潘俊帆，「以貝氏網路為基礎的個人差異化知識處理機制」，碩士論文，私立東海大學工業工程研究所，2001。
- [26].廖嘉雄，「以貝氏網路為基礎的知識處理機制」，碩士論文，私立東海大學工業工程研究所，2000。