Pertanika J. Soc. Sci. & Hum. 22 (3): 393 - 408 (2014)

PERTANIKA JOURNALS

SOCIAL SCIENCES & HUMANITIES

Journal homepage: http://www.pertanika.upm.edu.my/

# Identifying Factors Influencing Mathematical Problem Solving among Matriculation Students in Penang

**Ong, H. C.*** **and Lim, J. S.**

*School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM, Pulau Pinang, Malaysia.*

## ABSTRACT

Mathematics is recognized as an important subject in the school curriculum in Malaysia. It is a compulsory subject for many courses in matriculation, private colleges and universities. The purpose of this study is to identify the factors that influence the matriculation students in mathematical problem solving. Bayesian Network, a data mining technique, is used in this study to analyse the causal relationships. Bayesian network is a probabilistic graphical model which converts variables and their dependent relationships into nodes and arcs respectively. We compare the resultant networks using the different constraint and score based algorithms to identify the main factors affecting students in problem solving of mathematics. We found that students in Penang Matriculation College faced problem solving in mathematics owing to their problem with mathematical symbols. Hence, the students have no confidence in answering mathematics problems especially in questions related to their understanding of mathematical symbols.

*Keywords:* Bayesian Network, Learning Algorithms, Network Scores, Causal Relationship, Graphical Model, Mathematics Education, Data Mining.

## INTRODUCTION

Mathematics is recognized as an important subject in the school curriculum in Malaysia. In the Malaysian education system, students have to learn mathematics from kindergarten, right up to their matriculation studies. In private colleges and universities, mathematics is a compulsory subject in many courses. Students can apply for admission to matriculation courses which are coordinated and carried out by the Ministry of Education (MOE) (Hong *et al*, 2009). In the application, mathematics and additional mathematics are two important subjects that are considered for admission into the matriculation programme. Most of

ARTICLE INFO

*Article history:*
Received: 14 June 2012
Accepted: 2 August 2012

*E-mail addresses*:
hcong@cs.usm.my (Ong, H. C.),
luckjs80@yahoo.com (Lim, J. S.)
* Corresponding author

ISSN: 0128-7702 

the selected applicants have good grades in these two subjects. However, the majority of the matriculation students under the 1 year programme still face difficulties on problem solving in mathematics. This study uses a questionnaire survey to gather information needed from the students in Penang Matriculation College session 2010/2011. We then use Bayesian Network to analyze the causal relationships of the students on problem solving in Mathematics. The Objectives of this study are to explore the mathematics problems faced by students in the Penang Matriculation College and to use Bayesian Network to identify the most significant mathematics problem faced by students in the matriculation programme. Bayesian network is used in this study in place of other statistical methods like regression because it does not fix or assume the variables to be dependent or independent. Instead, the structural learning in Bayesian network explores the structural dependencies among the variables.

A Bayesian Network is a probabilistic graphical model that encodes variables and their dependent relationships into nodes and arcs respectively (Heckerman, 1995; Pearl, 1986). In general, Bayesian Network can be defined as follows: Assume that $S = \{G, \theta\}$ be a joint probability distribution of a set of $n$ random variables $X = \{X_1, X_2, \ldots, X_n\}$ and is specified by a directed acyclic graph $G$ with a set of conditional probability functions parameterized by $\theta$ (Pearl, 1988). According to Cao and Fang (2010), the Bayesian Network structure, $G$, encodes the probabilistic dependencies in the data and the presence of an edge between two variables means that there exists a direct dependency between them. This set $S$ contains the parameter $\theta_{x_i|\pi_i} = P_S\left(x_i \mid \pi_i\right)$ for each realization $x_i$ of $X_i$ conditioned on $\pi_i$, the parents of $x_i$ in $G$. Thus, $S$ can be defined as a unique joint probability distribution over $X$, written as $P_S\left(X_1, X_2, ..., X_n\right) = \prod_{i=1}^{n} P_S\left(X_i \mid \pi_i\right) = \prod_{i=1}^{n} \theta_{X_i|\pi_i}$, where $\pi_i$ represents the causes (parents) of variable $X_i$. Bayesian Network is an approach to detect causal structures in data (Pearl, 2000). We know that Bayesian Network is a graphical representation of a probabilistic distribution on a set of random variables. Tchangani (2002) stated that Bayesian Network has a graphical representation of causality relationship between a cause and its effect. The nodes are linked by directed arcs that create a Directed Acyclic Graph (DAG) and the DAG shows no route or path from one node connecting back to itself or else it will be a cyclic graph. However, the arcs represent the conditional independent relationships between the nodes. Assume that an arc from node $R$ to node $Q$ shows that the probability specification for node $Q$ is directly dependent on the values in node R. In this case, R is called *a parent* of node $Q$ and node $Q$ is called a *child* of node R. Figure 1 shows the relationship between the nodes $R$ and $Q$.

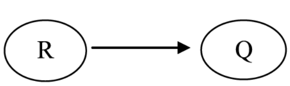


Fig.1: Connection from node $R$ to $Q$

The direction of the arrow shows the state of information of the decision maker, that is, whether the decision maker is capable of expressing the probability as P ($Q \mid R$).

## LITERATURE REVIEW

In the early development of Bayesian Network, it is used to solve problems in computational complexity and independence assumptions (Ni *et al*., 2010). According to Tchangani (2002), Bayesian Networks derive from convergence of statistical methods that allow one to go from information (data) to knowledge (such as probability laws and relationship between variables) and Artificial Intelligence (AI) that allow computers to deal with knowledge. Pearl (1988) and Jensen (1996) both agree that Bayesian Networks (BNs) are among the leading technologies to describe and derive conditional independence relationship among the random variables. Bayesian Network has become a powerful tool for causal relationship modelling and probabilistic reasoning (Tang *et al*., 2010) because researchers use Bayesian Network to handle problems with much greater complexity. It has become advantageous in a variety of areas including medicine (Gevaert *et al*, 2006), environmental protection (Henriksen and Barlebo, 2007) and financial risk management (Neil *et al*., 2005).

In Malaysia as in most countries, mathematics is a compulsory subject for students. Aziz (2005) stated that in learning mathematics, students always encounter problems involving calculations, understanding of concepts, principles and mathematical relationship with the others subjects. Norah *et al*. (2009) also claimed that learning of mathematics is a dynamic and complex process due to the interaction between previously acquired levels of understanding, conceptualization and incorporating of new materials. However, mathematics is more challenging for students. In matriculation, students often complained that mathematics is hard to learn and difficult to relate to in their studies. According to Haron *et al*. (2000), mathematics is the most difficult subject to understand among the students in the matriculation programme of Universiti Kebangsaan Malaysia (UKM). In Mahmud (2003), the main reason for secondary school students' difficulties in solving mathematical problems was an inability to understand the problem. Hong (2004) found that students had problems in solving non-routine mathematical problems even though they could pass their mathematics examinations. Aziz (2005) claimed that mathematics is difficult to learn because the concept in mathematics is abstract and hard to understand. Irvin and Norton (2007) claimed that students' poor attitudes toward mathematics cause them to perceive mathematics as a dry and static subject, abstract and only involved calculation.

In this study, problem solving in mathematics is an issue we focus on. In the start of the 21st century, the Ministry of Education (MOE) in Malaysia emphasizes that problem solving is one of the various aspects in teaching and learning when implementing the revised curriculum (MOE, 2001). Chapman (2005) stated that problem solving is important as a method for learning and teaching mathematics. Zakaria *et al*. (2009) said that solving a problem is a task in which an individual uses his/her existing knowledge, skills and understanding to address an unfamiliar situation. A problem solver also needs a rich, connected understanding of mathematics, ability to see patterns of similarities and association, skills to carry out the solution and finally, check that the results make sense in context of the problem (Burkhardt and Bell, 2007).

## METHODOLOGY

The sample involves 1312 students from Penang Matriculation College in the academic session 2010/2011. The respondents are of the same age and similar educational background where all have passed their PSPM 1 (Peperiksaan Semester 1 Program Matrikulasi) semester 1 examination. In addition, the respondents will sit for their PSPM 2 soon. This study used a questionnaire that is similar to Chong (2006) but is modified to cater for matriculation students. This questionnaire consists of twelve items. All the twelve items are given in five Likert scales, with 1 denoting "Strongly Disagree", 2 denoting "Disagree", 3 denoting "Neutral", 4 denoting "Agree" and 5 denoting "Strongly Agree". This instrument is designed to see a causal relationship between the items and all items are related to problem solving in mathematics.

This instrument contains 12 questions. They are from Q1 until Q12. All the questionnaires from Q1 to Q12 are available in Appendix 1. The questionnaires are distributed to the students during their lectures.

### *Structural Learning in Bayesian Network*

Structure Learning Algorithms which include Scored-based (Singh and Valtorta, 1995; Margaritis, 2003) and constraint-based (Cooper, 1997; Margaritis, 2003) are two categories of structure learning algorithms for Bayesian Network. The score-based method uses a score metric that measures how a structure reflects the data and finds a Bayesian Network structure with the highest score (Na and Yang, 2010). However, the DAG in constraint-based method is based on a set of conditional independent statements and is recognized from some prior knowledge or on some calculation from the data (Margaritis, 2003).

### *Analysis in the Bayesian Network*

In this study, we use both score based methods and constraint based methods to determine the major causes for students to be poor in solving mathematics problems. In learning a large system, heuristic algorithms such as Hill – Climbing (*HC*) are commonly used in practice (Kojima *et al*., 2010). Kojima *et al*. (2010) also claimed that the Hill – Climbing algorithm is used to find the local optima and upgraded versions of this algorithms lead to improving the score and structure of the results. Daly and Shen (2007) stated that the optimised implementation uses score caching, score decomposability and finally score equivalence. These scores will reduce the number of duplicated tests. Grow –Shrink (GS) algorithm consists of two phases which are a grow phase and a shrink phase. The GS algorithm was proposed by Margaritis (2003). In Tsamardinos *et al*. (2003), Incremental Association Markov Blanket ( IAMB ) algorithm is based on the Markov Blanket detection algorithm which consists of two phases: a forward phase and a backward phase. Interleaved Incremental Association Markov Blanket (Inter- *IAMB)* is another variant of IAMB. It has two phases: growing phase and shrinking phase. It used a forward stepwise selection which avoids false positives in the Markov Blanket. (Tsamardinos *et al*., 2003 ; Ge *et al*., 2010)**.** Fast Incremental Association Markov Blanket (Fast- *IAMB)* also contains two phases: growing phase and shrinking phase (Yaramakala and Margaritis, 2005). It is similar to GS and IAMB. An algorithm that is called max – min hill climbing (MMHC) proposed by Tsamardinos *et al*. (2006), combined an independence test (IT) approach with a score based strategy where an undirected graph is constructed or built depending on an IT approach and a constrained greedy hill climbing search which returns a local optimum of the score function. Restricted Maximization (RSMAX2) is a more general implementation of the Max-Min Hill-Climbing, which can use any combination of constraint-based and score-based algorithms (Scutari, 2010). Thus, HC and RSMAX2 used the scored based method while the other learning algorithms used constraint based methods. A score based Bayesian network structure search is used in Tamada *et al*. (2011) to find the DAG structure fitted to the observed data and the score function is used to measure the fitness of the structure to the given data. Ge *et al*. (2010) stated that a score function Score (*G*, *D*) for learning a Bayesian network structure is decomposable. It can be expressed as a sum of local scores. Score (G, *D*) = $\sum_{i=1}^{m} S(D_i, D_{Gi})$ where *G* is a directed acyclic graph (DAG) and *D* is a certain data set. There are several scores proposed for learning Bayesian networks such as the Bayesian Dirichlet equivalent or Bde (Heckerman *et al*., 1995), the Bayesian Information Criterion or BIC (Schwarz, 1978), the Akaike Information Criterion or AIC (Akaike, 1974) and the greedy heuristic algorithm or K2 (Cooper and Herskovits, 1992). We calculate the score results based on networks obtained from the seven learning algorithms, which are Hill-Climbing (HC) , Grow- Shrink (GS), Incremental Association Markov Blanket (IAMB), Fast Incremental Association Markov Blanket (Fast-IAMB), Interleaved Incremental Association Markov Blanket (Inter-IAMB), Max - Min Hill Climbing (MMHC) and Restricted Maximization (RSMAX2). These score functions are used to estimate the network fit for these algorithms. Score-based methods produce a series of candidate Bayesian networks from the learning algorithms; calculate a score for each candidate and return a candidate of highest score (Jensen, 2009). Akaike Information Criterion or AIC was developed by Akaike (1977). Akaike (1973) used the AIC to select the model that minimizes the negative likelihood penalized by the number of parameters as specified in the equation (1).

$$AIC = -2 \log p(L) + 2p \quad (1)$$

where $L$ refers to the likelihood under the fitted model and $p$ is the number of parameters in the model. It is used to find the approximate model to the unknown true data (Acquah, 2010). Another information criterion that is widely used is BIC or Bayesian Information Criterion. BIC is derived within a Bayesian framework as an estimate of the Bayes factor for two competing models (Schwarz, 1978; Jensen, 2009). The score for the BIC is defined as

$$BIC = -2 \log p(L) + \log n \quad (2)$$

where $n$ is a sample size. The difference between AIC and BIC is based on the second term which is the sample size (Acquah, 2010). Heckerman *et al*. (1995) developed the Bde or Bayesian Dirichlet Equivalent score. This score uses Bayesian analysis to evaluate and estimate a given dataset network. The idea of Bde is dependent on the BD (Bayesian Dirichlet) metric which is developed by Cooper and Herskovits (1992). The Dirichlet distribution is a multivariate distribution to describe the conditional probability of each variable in the network. The algorithm of K2 score is another posterior density which is proposed by Cooper and Herskovits (1992). The K2- like greedy search method will incrementally add a node to a parent set and find the best parent set to maximize the joint probability of the structure and the database (Yang *et al*., 2006). The log-likelihood (loglik) score is equivalent to the entropy measure used in Weka (Witten and Frank, 2005). The maximized likelihood $P(D|G)$ decomposed by the network structure and for the decomposable scores is the complexity penalty.

## RESULTS AND DISCUSSION

Similar to Ge *et al*. (2010), the *bnlearn* package (R Team 2009) in R is used to run the structural learning algorithms. From the structural learning algorithms, there are seven different networks outcomes which are noncyclical. The arcs show direct dependent relationships between the connecting variables. However, the existence of conditional independence relationships is indicated by the absence of arcs (Ge *et al*. 2010). These diagrams also represent the logical cause and effect between the variables. Table 1 shows the numbers of edges and arcs for each pair of the learned networks. The "edges" represent the number of common links or edges (in either direction) for each learning networks structure. However, the "arcs" represent the number of common directed arcs between the nodes in these learned networks. Table 1 also shows the number of common links and arcs that are obtained in each network in the diagonal section. From Table 1 a number of nodes that were constructed or built are the same. The nodes with the common arcs for all models in these learned networks represent a strong relationship in between the connections in these nodes. The edges with strong relationships are Q1 to Q2, Q4 to Q5, Q5 to Q6 and finally Q11 to Q12. Besides that, the other nodes that show the weak relationships in between them are Q3, Q7, Q8, Q9 and Q10. Figure 2 shows all the learned networks of the various algorithms. The common arcs are shown in Figure 3. These common arcs show that there is only one way direction to the consecutive nodes in all these learned networks.

For instance, node Q5 links to node Q6. The connection of edges from Q5 to Q6 can be interpreted as students being sure of which method to be used when encountering a long mathematics question because they do not know what information is needed to handle the mathematics question. This happens because they do not understand the question and fail to transform their idea into mathematics symbol. Figure 3 shows the directly connected nodes. These links between the nodes represent common edges to all of the learned networks. Following this, we run again these seven learned network algorithms and set the common edges using Figure 3 as a white list for each learned networks. Then, we obtained the result for all the seven new learned networks (after white list) in Figure 4. We also show the final result on the HC network in Figure 5.

By running all the algorithms, the result of scores for the seven algorithms are shown in Table 2. The obtained results are important for comparing the network from the algorithms. In this study, network scores are used because they select which network fitted the data best. Based on the results shown in Table 2, we highlighted the highest scores for the networks. Following the white list of all the learned networks and having set the arcs, we found the Hill – Climbing ( HC ) algorithm as having the best result for this study from Table 2. The arc strength is used to evaluate the strength for all the edges. Each edge will show the highest and the lowest score of the strength. The arc strength is used to measure the strength of the probabilistic relationships expressed by the arcs of a Bayesian network and it uses model averaging to build a network containing only the significant arcs (Scutari, 2010).

In Figure 5, the thicker arcs represent the stronger relationships between the nodes.

TABLE 1
Number of common edges/arcs between each pair of the learned networks

| | Hc | Gs | Iamb | Fast. iamb | Inter. iamb | Mmhc | Rsmax2 |
|---|---|---|---|---|---|---|---|
| Hc | 11/11 | 8/1 | 6/1 | 6/2 | 6/1 | 5/3 | 8/8 |
| Gs | - | 13/7 | 6/2 | 6/2 | 6/2 | 5/0 | 11/1 |
| Iamb | - | - | 12/9 | 10/6 | 10/9 | 10/2 | 6/1 |
| Fast. iamb | - | - | - | 11/10 | 10/6 | 9/3 | 6/2 |
| Inter. iamb | - | - | - | - | 12/9 | 10/3 | 6/1 |
| Mmhc | - | - | - | - | - | 10/10 | 5/3 |
| Rsmax2 | - | - | - | - | - | - | 11/11 |

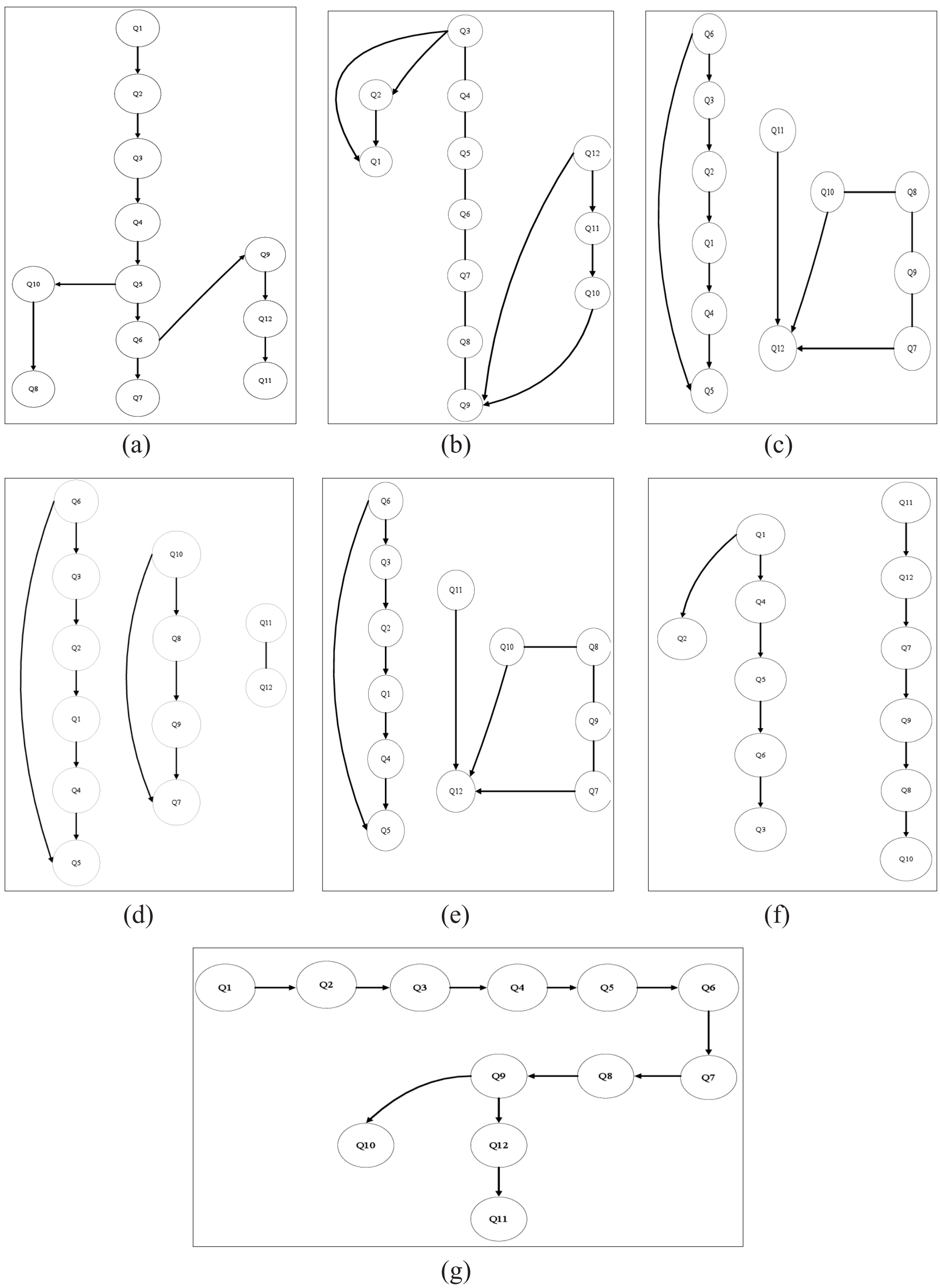


Fig.2: Network structures learned by selected algorithms. (a) Hill – Climbing; (b) Grow – Shrink; (c) Incremental Association Markov Blanket; (d) Fast Incremental Association Markov Blanket; (e) Interleaved Incremental Association Markov Blanket; (f) Max – Min Hill Climbing; (g) Restricted Maximization

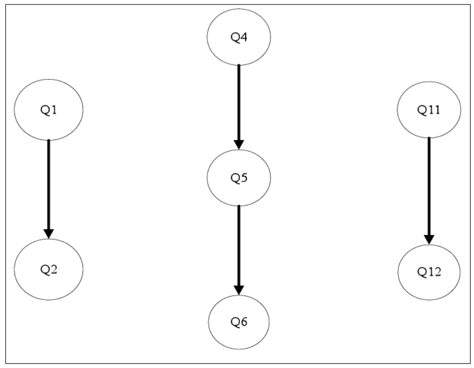


Fig.3: Common edges of all the learned networks

TABLE 2
The results of scores of all learned networks for each algorithm

| | Aic | Bic | Bde | Loglik | K2 |
|---|---|---|---|---|---|
| HC | **-20568.34** | **-21521.33** | **-21852.35** | -20200.34 | **-20893.12** |
| GS | -21336.91 | -24527.37 | -23366.60 | -20104.91 | -21565.67 |
| IAMB | -22943.70 | -31106.29 | -24796.8 | **-19791.70** | -22091.64 |
| Fast-IAMB | -20961.85 | -22287.75 | -22337.56 | -20449.85 | -21307.40 |
| Inter-IAMB | -22943.70 | -31106.29 | -24796.8 | **-19791.70** | -22091.64 |
| MMHC | -20758.08 | -21711.08 | -21989.83 | -20390.08 | -21079.28 |
| RSMAX 2 | -20883.56 | -21795.12 | -22048.56 | -20531.56 | -21188.50 |

These arcs also represent the highest values in arc strength compared with the others. However, the thin lines that are shown in the network are edges that represent the supplementary edges. Based on the Figure 5, we displayed the stronger relationship and the highest value of arc strength in the final result of the learned network in Table 3.

TABLE 3
The stronger relationship between the nodes and the arc strength in the final result learned network

| Edges | Arc strength |
|---|---|
| Q2 to Q12 | 198.92559 |
| Q4 to Q6 | 186.66152 |
| Q11 to Q12 | 113.60172 |
| Q5 to Q6 | 102.7143 |

Table 3 is from the HC algorithm which gives the best scores among the seven learned networks except the log-likelihood scores. IAMB and Inter- IAMB both obtained the same highest score compared with the others. Therefore, the result from HC algorithm (from Figure 4(a) and in Figure 5) is the learned network from which we select the final result of this study. In Figure 5, the arc from node 2 to node 12 represents the strongest relationship among the nodes. Based on the questionnaire, due to students' abilities in solving the mathematics questions, they have difficulties with the complicated mathematical symbols and this causes students to have no confidence in coming

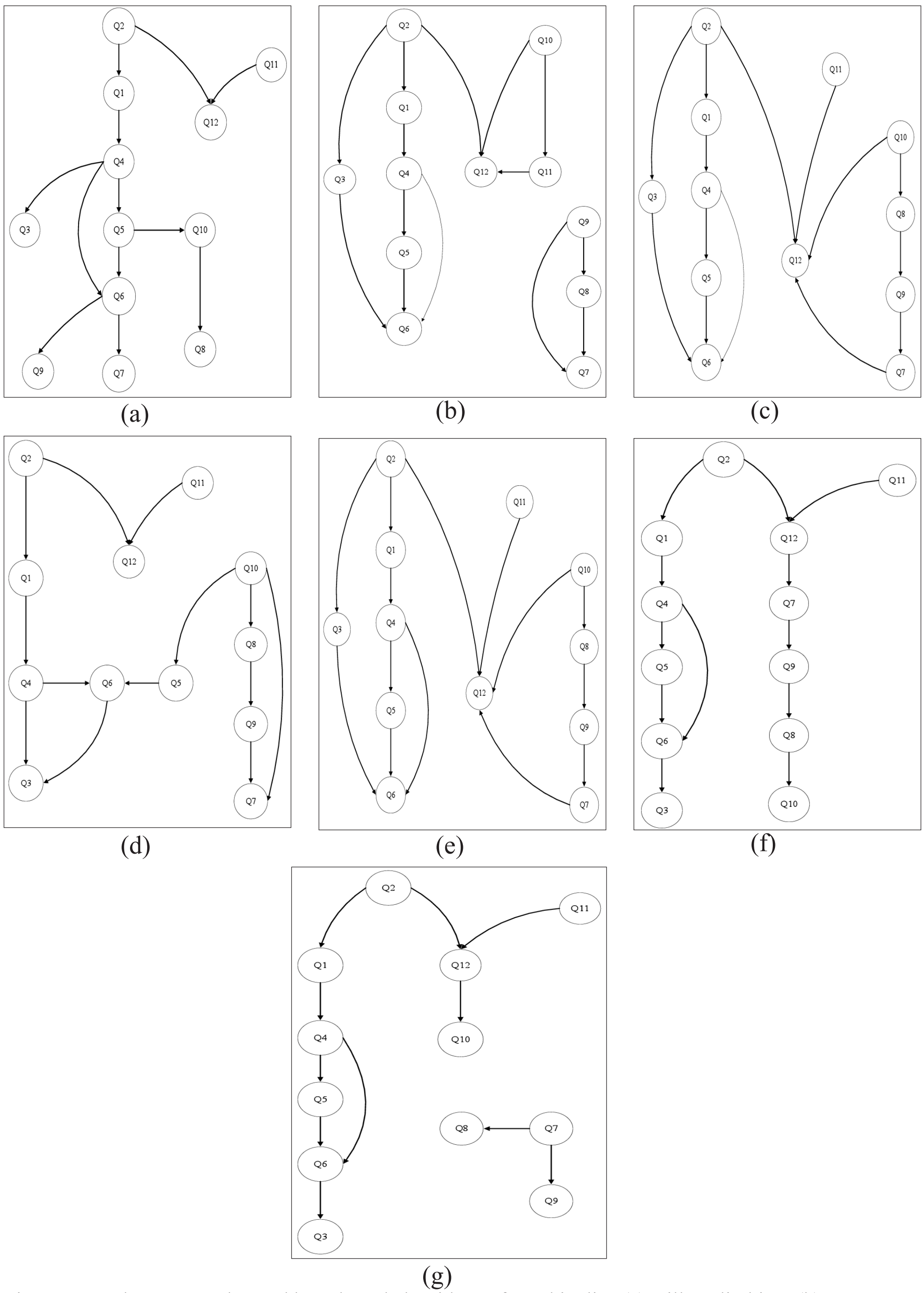


Fig.4: Network structures learned by selected algorithms after white list. (a) Hill – Climbing; (b) Grow – Shrink; (c) Incremental Association Markov Blanket; (d) Fast Incremental Association Markov Blanket; (e) Interleaved Incremental Association Markov Blanket; (f) Max – Min Hill Climbing; (g) Restricted Maximization

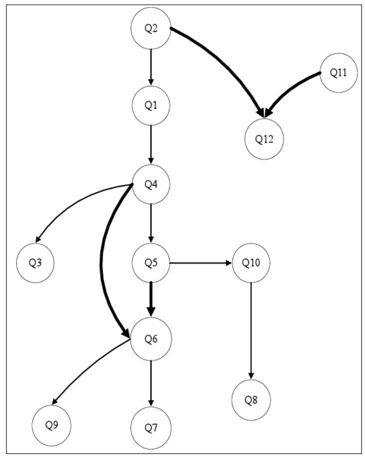


Fig.5: The final result of the score learned network using the Hill – Climbing algorithm.

out with a neat and complete solution in solving the mathematics question. The final result of learned network shows that this is a major factor that causes students to be poor in solving mathematics problems. Similarly in Kinzel (1999), students have difficulties in understanding and interpreting the symbolic notation used in algebra. Capraro and Joffrion (2006) claimed that middle-school students often demonstrated much stronger skills in solving formal and informal problems that require algebraic reasoning than in symbolizing equations. They also indicate those students' abilities to solve simple word problems with arithmetic and should be connected to the formal algebraic symbolic notation.

Furthermore, the arc from node 4 to node 6 shows the second major problem faced by students in this study. We found that students who lack understanding of the mathematics question requirement, failed to transform the question needed into mathematical symbols which causes them to be not sure of which method to use if they are faced with a long mathematics question. Also in Ilany and Margolin (2010), language of symbols, concepts, definition and theorems are considered mathematical language. They also mentioned that the mathematical language needs to be learned and it cannot be developed naturally like a child's natural language. The arc from Node 11 to node 12 gives the third highest score for the strength in Table 3. Students always make careless mistakes in the process of calculation during their attempt to solve the mathematics questions. The mistakes that they make will lead them to be weak and poor in coming out with a complete solution in solving mathematics problem. Students are unable to write the appropriate solution for the mathematics question given because they do not plan well and organize in solving mathematics problem. From Montague (1988), some students with learning disabilities may have learned and organized correct strategies and conceptual knowledge to solve problems, but then fail to carry out them as is required.

## CONCLUSION

The major problem solving in mathematics that are faced by students in Penang Matriculation College is due to their understanding of mathematical symbols that influence their abilities in solving mathematics problems. From the Bayesian Network, this score is the highest in the final result of learned network. Owing to the complicated and difficult mathematical symbols, the students are unable to perform their solution well or in other words; they cannot solve the mathematics question with a neat and complete solution. Furthermore, our study also revealed that students in Penang Matriculation College are confused with the method to be used when they are faced with long mathematics questions. We can also conclude that the students are afraid of the complicated mathematics symbols and do not really understand the questions needed when they try to solve the questions. This study also revealed that students are quite weak in transforming the information from mathematics questions into the mathematics language.

Bayesian Network is a powerful tool to trace problems in many areas such as in the industry and data mining. In our study, we use Bayesian network algorithms in the mathematics learning situation to identify the problems which arise. The results obtained show that the students in Penang Matriculation College have difficulty in understanding complicated mathematics symbol which causes them to be not confident in coming out with a neat and complete mathematical solution. Having identified mathematics symbols as the root cause of the problem in mathematical problem solving, future and subsequent work can be carried to help students based on their understanding of the various types of mathematics symbols.

## ACKNOWLEDGEMENTS

This work was supported in part by U.S.M. Fundamental Research Grant Scheme (FRGS) No. 203/PMATHS/6711319.

## REFERENCES

Acquah, H. (2010). Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship. *Journal of Development and Agricultural Economics*. *2*(1), pp. 1 – 6.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control.19* (6), pp. 716 –723.

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In: *B.N. Petrov and F. Csaki (eds.) 2nd International Symposium on Information Theory*: pp. 267-281. Budapest: Akademiai Kiado.

Akaike, H. (1977). On entropy miximisation principle. In: *Proc. Symp. On Application of Statistics ed. P.R Krishnaiah*, pp. 27 – 47. Amsterdam, the Netherlands.

Aziz, N. (2005). Students' Perception on Teaching And Learning Mathematics In English. Buletin Pendidikan Sains dan Matematik Johor, *14*(1). ISSN 0128-4290.

Burkhardt, H., & Bell, A. (2007). Problem solving in the United Kingdom.*ZDM Mathematics Education*. *39*, pp. 395 – 403.

Cao, W., & Fang, X. ( 2010). An Improved Method for Bayesian Network Structure Learning. In: *Sixth International Conference on Natural Computation(ICNC 2010)*, 3133 – 3137.

Capraro, M. M., & Joffrion, H. (2006). Algebraic Equations: Can Middle-school Students Meaningfully Translate from Words to Mathematical Symbols?. *Reading Psychology*. *27* , pp. 147 – 164.

Chapman, O. (2005). Constructing Pedagogical Knowledge of Problem Solving: Preservice Mathematics Teachers. In: *Proceeding of the 29th Conference of the International Group for the Psychology of Mathematics Education*, *2*, pp. 225 – 232.

Chong, C. L. (2006). Masalah Penyelesaian Soalan Matematik Berayat Di Sekolah Kebangsaan Batu Lima, Gurun, Kedah. M.Sc. Thesis, USM.

Cooper, G. F. (1997). A Simple Constraint-Based Algorithm for Efficiently Mining Observational Databases for Causal Relationships. *Data Mining and Knowledge Discovery*, 1, pp. 2203– 2224.

Cooper, G. F., & Herskovits, E. (1992). A bayesian method for the induction of probabilistic networks from data. *Machine Learning*, *9*, 309 – 347.

Daly, R., & Shen, Q. (2007). Methods to accelerate the learning of Bayesian network structures. In: *Proceedings of the 2007 UK Workshop on Computational Intelligence*, Imperial College, London (2007).

Gevaert, O., DeSmet, F., Kirk, E. Van Calster, B., Bourne, T, Van Huffel, S., Morean, Y., Timmerman, D., De Moor, B., & Condous, G. (2006). Predicting the outcome of pregnancies of unknown location: Bayesian networks with expert prior information compared to logistic regression. *Human Reproduction*. *21* (7), pp.1824–1831.

Ge, Y., Li, C., & Yin, Q. (2010). Study on Factors of Floating Women's Income in Jiangsu Province Based on Bayesian Networks. *Advances in Neural Network Research and Applications*, Lecture Notes in Electrical Engineering 2010. *67*(9), pp. 819–827.

Haron, Z., Amin, L., & Ashaari @Shaari, N. (2000). Meninjau Masalah Pembelajaran Pelajar Matrikulasi UKM di Kampus Bangi. Strategi Pengajaran Kepelbagaian Pelajar. In: *ProsidingSeminar PendidikanKebangsaan. FakultiPendidikan* UKM, pp.252-263.

Heckerman, D. (1995). A tutorial on learning with Bayesian networks.*Technical Report MSR-TR-95-06*, Microsoft Research, Redmond, Washington.

Heckerman, D., Geiger, D., & Chickering, D.M. (1995). Learning Bayesian networks: the combination of knowledge and statistical data. *Machine Learning*, *20* (3), pp.197–243.

Henriksen, H. J., & Barlebo, H. C. (2007).Reflections on the use of Bayesian belief networks for adaptive management. *Journal of Environmental Management*, *88*(4), 1025–1036.

Hong, K. S. (2004). Student teachers ability to solve non-routine calculus problems. In: *Proceedings of Seminar Kebangsaan Pendidikan Guru 2004(CD Rom).*UPSI-JPG, Tanjung Malim.

Hong, K. S., Ting, L. N., & Hasbee, H. U.(2009). Status of mathematics teaching and learning in Malaysia. *International Journal of Mathematical Education in Science and Technology*, *40*(1), pp.59-72.

Ilany, B-S., & Margolin, B. (2010). Language and Mathematics: Bridging between Natural Language and Mathematical Language in Solving Problems in Mathematics. *Creative Education*, *1*(3), 138 – 148.

Irvin, J., & Norton, S. (2007). Developing Positive Attitudes Towards Algebra. In: *Proceedings of 30th Annual Conference of the Mathematics Education Research Group of Australia*. J. Watson and K. Beswick (Eds), © MERGA Inc. 2007. 561 – 570.

Jensen, F. V. (1996). *An Introduction to Bayesian Networks*. Springer Verlag, New York, N.Y.

Jensen, F. V. (2009). Bayesian Networks. In: *Wiley Interdisciplinary Reviews: Computational Statistics*. *1*(3), 307 – 315.

Kinzel, M. T. (1999). Understanding algebraic notation from the students' perspective. *Mathematics Teacher*, *92*, 436 – 442.

Kojima, K., Perrier, E., Imoto, S., & Miyano, S. (2010). Optimal Search on Clustered Structure Constraint for Learning Bayesian Network Structure. *Journal of Machine Learning Research*. *11*, 285 – 310.

Mahmud, M. (2003). The use of problem solving heuristics and their relationship to achievement. *Journal of Technical Education*, *2*(1). Kuala Lumpur: Department of Technical Education, Ministry of Education, Malaysia.

Margaritis, D. (2003). Learning Bayesian Network Model Structure from Data.Ph.D.thesis, School of Computer Science, Carnegie-Mellon University, Pittsburgh, PA.

Ministry of Education. (2001). *Modul I: Pendekatan Pengajaran dan Pembelajaran Matematik KBSM*. [Modul I: Mathematics teaching and learning approach] Curriculum Development Centre, Ministry of Education, Malaysia.

Montague, M. (1988). Strategy Instruction and Mathematical Problem-Solving. *Reading, Writing and Learning Disabilities*, *4*, pp.275 – 290.

Na, Y., & Yang, J. ( 2010 ). Distributed Bayesian Network Structure Learning. *International Symposium on Industrial Electronics (ISIE)*, 2010 IEEE: 1607-1611.

Neil, M., Fenton, N., & Tailor, M. (2005). Using Bayesian networks to model expected and unexpected operational losses. Risk Anal *25*, pp.963–972.

Ni, Z. F., Philips, L. D., & Hanna, G. B. (2010). The Use of Bayesian Networks in Decision-Making. In *Surgical Research and Methodology*, pp.351 – 359, London: Springer – Verlag Berlin Heidelberg.

Norah, T., Azmi, S., & Ahmad, R. H. (2009). Evaluation of Student's Attitude Towards Learning of Busines Mathematics Subject: A Case Study. *Discovering Mathematics*. *31*(1), 29 – 35.

Pearl, J. (1986). Fusion, propagation and structuring in belief networks. *Uncertainty in Artificial Intelligence,* 357 – 370.

Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, SanMateo, CA.

Pearl, J. (2000). *Causality, models, reasoning, and inference*. Cambridge University press.

R Development Core Team. (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.

Scutari, M. (2010). Learning Bayesian network with the bnlearn R package. *Journal of Statistical Software, 35*(3), 1-22.

Singh, M and Valtorta, M. (1995). Construction of Bayesian network structures from data: a brief survey and an efficient algorithm. *International Journal of Approximate Reasoning*. pp.111– 131.

Tamada, Y., Imoto, S., & Miyano, S. (2011). Parallel Algorithm for Learning Optimal Bayesian Network Structure. *Journal of Machine Learning Research*. *12,* pp.2437-2459.

Tang, Y., Cooper, K., Cangussu, J., Tian, K., & Wu, Y. (2010). Towards Effective Improvement of the Bayesian Belief Network Structure Learning. In: *International Conference On Intelligence and Securing Informatics (ISI)*, 2010, IEEE, p.174.

Tchangani, A. P. (2002). Wireless Local Loop Media Modelling using Bayesian Networks. *Systems Analysis Modelling Simulation*, *42*(1), 33 – 47.

Tsamardinos, I., Aliferis, C. F., & Statnikov, A. (2003). Algorithms for large scale Markov blanket discovery. In: *Proceedings of the Sixteenth International Florida Artificial Intelligence Research Society Conference*, 376–381. AAAI Press, Menlo Park.

Tsamardinos, I., Brown, L. E., and Aliferis, C. F. (2006). The max-min hill-climbing Bayesian network structure learning algorithm. *Machine Learning*, *65*, 31–78.

Witten, I. H., & Frank, E. (2005). Data Mining: Practical Machine Learning Tools and Techniques, 2nd Edn. Morgan Kaufmann, San Francisco.

Yang, Y.L., Wu, Y., & Liu, S. Y. (2006). The Constructing of Bayesian Networks Based on Intelligent Optimization. In: *Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications (ISDA'06)*. *1*, pp.371- 376.

Yaramakala, S., & Margaritis, D. (2005). Speculative Markov Blanket Discovery for Optimal Feature Selection. In: *Proceeding of the Fifth IEEE International Conference on Data Mining(ICDM), 2005*., pp.809-812, Houston, Texas.

Zakaria, E., Yazid, Z., & Ahamd, S. (2009). Exploring Matriculation Students' Matacognitive Awareness and Achievement in a Mathematics Course. *The International Journal of Learning*. *16* (2), pp. 333 – 347.

## APPENDIX 1

The Questionnaire for Mathematics Problems Faced By Matriculation Students

This questionnaire is designed to collect data regarding mathematics problems faced by matriculation students.

You are required to answer all the questions sincerely. There is no right or wrong answers. Please circle your preference.

Guideline: 1 = Strongly Disagree
2 = Disagree
3 = Neutral
4 = Agree
5 = Strongly Agree

| | |
|---|---|
| 1. I am lacking in ability to solve the question given because I do not understand the words/ phrases in mathematics. | 1 – 2 – 3 – 4 – 5 |
| 2. Complicated mathematics symbols in solving mathematics question reduces my ability. | 1 – 2 – 3 – 4 – 5 |
| 3. Long and complicated / tough questions hinder me from solving the mathematics question. | 1 – 2 – 3 – 4 – 5 |
| 4. I do not understand the question requirement and fail to transform to mathematics symbol. | 1 – 2 – 3 – 4 – 5 |
| 5. I do not know/ am not sure what information is needed at tackling mathematics question. | 1 – 2 – 3 – 4 – 5 |
| 6. I am not sure of which method to be used when faced with long mathematics question. | 1 – 2 – 3 – 4 – 5 |
| 7. I cannot change to an alternative method when I am stuck half way with the method used to solve question. | 1 – 2 – 3 – 4 – 5 |
| 8. I always make mistakes when solving mathematics question because I am not familiar with the basic operation of mathematics (+, - , x , ÷). | 1 – 2 – 3 – 4 – 5 |
| 9. I am nervous when faced with long mathematics question because I am unable to connect/ to link the theory that I have learned. | 1 – 2 – 3 – 4 – 5 |
| 10. I am always forget the symbol to be used to solve mathematics question. | 1 – 2 – 3 – 4 – 5 |
| 11. I am always careless in the process of calculation when solving mathematics questions. | 1 – 2 – 3 – 4 – 5 |
| 12. I have no confidence in coming out with a neat and complete mathematics solution. | 1 – 2 – 3 – 4 – 5 |