Paper: jc*--**-****:**

Analysis of Influence Factors for Learning Outcomes with Bayesian Network

Kazushi Okamoto

Department of Informatics, Graduate School of Informatics and Engineering, The University of Electro-Communications
E-mail: kazushi@uec.ac.jp
[Received 2018/02/26; accepted 2018/07/23]

Abstract. This study identifies and analyzes the influence factors for learning outcomes at a university with a Bayesian network. It is based on a fact-finding survey on university student life and learning. Suitable constraints and a score metric for the Bayesian network learning are determined via cross-validation, and the learning outcome variables are categorized into subsets according to six abilities: cooperativeness, expressiveness, foreign language, collecting and organizing information, logical thinking, and sociability. The learned network suggests that two to seven factors influence each ability. In addition, it is confirmed that the probability distributions of all most of the identified factors shift to high agreement/experience levels, as self-knowledge levels for the acquired abilities increase, i.e., positive effects exist for most factors for each identified ability.

Keywords: Bayesian network; learning outcome; causal relationship; conditional probability

1. Introduction

Identifying influence factors for the learning outcomes at a university is indispensable in terms of verifying the services and environments that a university should provide for the students' learning. Every university has an interest in this and deals with such verification as a part of its institutional research. In particular, surveys based on questionnaires, interviews, and the analysis of academic results are widely applied. On the other hand, with only one university's survey, it is difficult to identify and generalize the factors that influence the students' learning outcomes because the survey items and student characteristics vary among different universities. Analysis of national-level data is therefore important.

In Japan, some nationwide surveys have been performed for example, a nationwide university student survey by Department of University Management and Policy Studies, The University of Tokyo in 2007[1]; a fact-finding survey on university student life and learning by the Benesse Educational Research and Development Institute in 2008[2]; a survey on university student learning by the National Institute for Educational Policy Research in 2014[3]. The fact-finding survey on university student life and learning published the raw response data for free in 2014, and the quantity of responses for it is over 4,000. Previous research, which analyzes the raw response data secondly, has been published[4, 5]. The applied analytical methods are based on interval or ratio scales such as factor analysis and covariance structure analysis. Such applications, however, are unsuitable, since the questionnaires in the survey are presented to subjects with response alternatives consisting of nominal or ordinal scales.

This study performs a secondary analysis of the fact-finding survey on university student life and learning with a Bayesian network[6–8]. We analyze and identify the factors that influence learning outcomes at a university. The Bayesian network is an approach to estimate and visualize the causal relationships among variables with a directed acyclic graph (DAG) and conditional probability tables (CPTs). It can treat both discrete and continuous variables. In addition, assumptions about the models and latent factors are unnecessary. We, thus, consider that an analysis based on a Bayesian network is effective for questionnaires consisting of response alternatives with nominal or order scales and dealing with many variables.

In this study, four question groups in the survey are used: (i) learning outcomes at the university; (ii) opinions about education in the university; (iii) attitude towards learning activities; (iv) class experience. We use the questions within (i) as response variables and categorize the subset according to the six abilities: cooperativeness; expressiveness; foreign language; collecting and organizing information; logical thinking; sociability. The questions within (ii), (iii), and (iv) are used as explanatory variables, and we identify and discuss how the explanatory variables influence the response variables.

In a Bayesian network, a suitable directed acyclic graph is obtained from the observed data. This process, called network learning, requires hyper-parameters such as a learning algorithm and a constraint for the arcs. This study preliminarily determines suitable hyper-parameters for the network learning via cross-validation; then, the influence factors for the learning outcomes are identified and analyzed based on the DAG and CPTs determined with the hyper-parameters.

This paper is organized as follows: Section 2 explains the details of the used dataset; Section 3 describes the application of the Bayesian network to the questionnaire;

Section 4 determines the hyper-parameters for network learning; Section 5 analyzes and identifies factors influencing learning outcomes at a university by using the learned DAG and CPTs.

2. Details of Dataset Used

This study also analyzes the fact-finding survey on university student life and learning, performed in 2008 by the Benesse Educational Research and Development Institute[2]. The survey targets were first to fourth year university students aged 18 to 24. University students, satisfing the above conditions, were randomly sampled from approximately 800,000 people depending on gender and faculty ratios determined by the 2008 statistics of the Ministry of Education, Culture, Sports, Science and Technology of Japan. The survey obtained 4,070 (male: 2,439, female: 1,631) responses, and the year distribution is 1,017 first year, 1,013 second year, 1,017 third year, and 1,023 fourth year students. The raw data was published via the Social Science Japan Data Archive in June 2014.

The main research areas are as follows:

- **face sheet**: gender; year; university division; university location; major faculty; major academic field (humanities/science); etc.
- **pre-university matters**: junior high school and high school entrance exam experience; type of high school; high school location; efforts in high school; studies in high school; important points for the choice of university; efforts towards university entrance exam; motivation for going to university; etc.
- **campus life**: university satisfaction; active participation in campus life; days spent at university; participation in clubs or societies; part-time work; commute time; how time is spent outside university; income per month; etc.
- **university studies**: attendance rates for classes; student status at university; academic results; things learned at university life (results of studies); etc.
- **post-graduation**: career consideration; preparation for job hunting and career; views of society and work; relationship with parents; etc.

In the questionnaire, there are a total of 291 questions, and each question is classified into 42 groups. This study uses four question groups: learning outcomes at the university (Question Group 37); opinions about education in the university (Question Group 32); attitude towards learning activities (Question Group 33); class experience (Question Group 35). These four question groups are all questions related to class-work in that survey, under "university studies". The number of questions used is 83 (10 + 26 + 19 + 28). Details about the used question groups, question sentences, and alternative responses are shown in Appendix A.

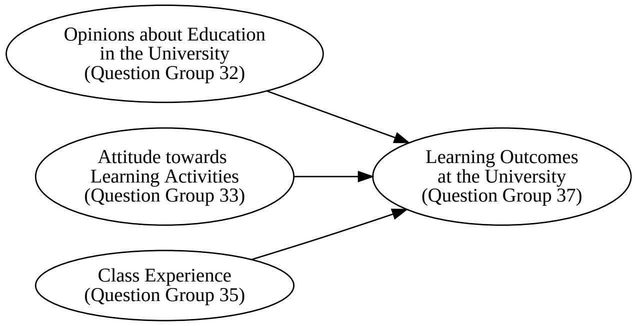

Fig. 1. Response and explanatory variables: variables within Question Groups 32, 33, and 35 explain the variables within Question Group 37.

With the four question groups, this study assumes that the learning outcomes at the university are influenced by the opinions about education in the university, the attitude towards learning activities, and the class experience, i.e., the response variables are the questions within Question Group 37 and the explanatory variables are the questions within Question Groups 32, 33, and 35. This assumption is shown in Fig.1: an arc is drawn from a cause to its effect, and it describes a causal relationship.

3. Data Analysis with Bayesian Network

A Bayesian network[6–8] is a kind of graphical model. It consists of DAG and CPTs. A DAG describes causal relationships among the variables: a node corresponds to a probability variable and an arc is drawn from a cause variable to a result variable. In this section, $X_1, X_2, \cdots, X_n$, X, Y, Z are defined as the probability variables.

3.1. Definition

A causal relationship from $X_1, X_2, \cdots, X_n$ to Y is written as the conditional probability $P(Y|X_1, X_2, \cdots, X_n)$. In a DAG, arcs are assigned from each of the $X_1, X_2, \cdots, X_n$ to Y. If the $X_1, X_2, \cdots, X_n$, and Y are discrete variables, $P(Y|X_1, X_2, \cdots, X_n)$ is generally calculated based on counting from the observed data and is stored as a CPT. Such counting and storage operations, however, have high computational costs, i.e., $r \times \prod_{i=1}^{n} r_i$, where r and r_i are the numbers of possible outcomes for Y and X_i, respectively.

On the other hand, conditional independence among Z and $X_1, X_2, \cdots, X_n$ is defined as

$$P(X_1, X_2, \cdots, X_n|Z) \rightleftharpoons \prod_{i=1}^{n} P(X_i|Z). \quad \ldots\ldots (1)$$

According to the definition of conditional probability,

$$P(Y|X_1, X_2, \cdots, X_n, Z) = P(Y|Z) \quad \ldots\ldots (2)$$

is established for Y, if Y and $X_1, X_2, \cdots, X_n$ are conditionally independent. This suggests that the conditional variables, which satisfy conditional independence, can be reduced.

In a DAG, conditional independence for two variables is described as the absence of an arc. For $X_1, X_2, \cdots, X_n$, a DAG corresponds to a calculation model of the joint probability $P(X_1, X_2, \cdots, X_n)$, since it satisfies

$$P(X_1, X_2, \cdots, X_n) = \prod_{i=1}^{n} P(X_i | X_{i-1}, \cdots, X_1) \quad . \quad . \quad (3)$$

from the chain rule in probability. The $P(X_i|X_{i-1}, \cdots, X_1)$ is reduced based on conditional independence.

3.2. Network Learning

In a Bayesian network, a DAG is automatically identified by an optimization algorithm or is manually given (both are possible). In the former, the best DAG is found from the observed data. This is called network learning. Network learning is a kind of combinatorial optimization problem, and its complexity is NP-complete[9]. Current proposed learning algorithms are roughly divided into score-based learning and constraint-based learning. Score-based learning consists of two parts: definition of an objective function for how well a DAG fits the observed data; searching over the space of possible DAGs to maximize the objective function. On the other hand, constraint-based learning consists of Markov blanket learning, neighbor identification with conditional independence test, and identification of arc directions.

3.3. Application to Data Analysis

In the application of a Bayesian network to data analysis, the network is able to identify the explanatory variables' effect on a response variable from a learned DAG. In addition, according to Bayes' theorem, $P(Y|X)$ can be inverted as

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

and the likelihood of a causal variable X can be estimated from a result variable Y. This study identifies causal variables from learning outcome questions based on their likelihood.

The Bayesian network is applicable to discrete variables, unlike analytical methods based on interval or ratio scales such as factor analysis and principal component analysis. Furthermore, correlations among explanatory variables are acceptable, unlike for linear regression analysis. Structural Equation Modeling (SEM) is a typical analysis approach for causal relationships. In the usage of SEM, an assumed model should be given by user. A Bayesian network, however, can be find the suitable model (DAG) from the observed data automatically; moreover, the user can also give a model or revise a found model manually.

3.4. Related Works

In terms of higher education, there has been much research that use Bayesian networks to analyze the collected data and to construct prediction models. Pumpuang *et al.*[10] and Sharabiani *et al.*[11] apply Bayesian networks to understand the relationships between courses and to construct prediction models. The datasets used, which consist of the students' GPAs and course grades, were collected at a private university in the Thailand and University of Illinois at Chicago, respectively. Fernández *et al.*[12] perform a case study with a dataset that consists of information and statistics on courses taught at the University of Almeria. In addition, they develop a Web-based advice system based on a learned DAG and probability tables. Millán *et al.*[13] give a tutorial and discuss how to model students using a Bayesian network. Grubišić *et al.*[14] propose a Bayesian network-based student modeling method with ontologies, and the prior probabilities are determined based on knowledge test results. Xenos[15] and García *et al.*[16] apply a Bayesian network to datasets collected from the Web-based education system. Watanabe *et al.*[17] validate educational effectiveness with a Bayesian network with regard to how a career design course affects to the students' self-understanding, communication skills, work considerations, and views in the future.

The previous studies focus on the construction of prediction models or identification of factors in terms of specific courses/majors or ability; however, this study covers all learning activities at a university. Unlike the previous studies, the dataset used in this study is not collected from a specific university or institute. In addition, the target students' majors not only included science and engineering but also the humanities.

4. Modeling with Bayesian Network

In Bayesian network learning, constraints for the existence and direction of the arcs for each node pair can be set, and many learning algorithms have thus been proposed. This section determines suitable constraints for the arcs and a score metric for the network learning that maximize the prediction accuracy from the learning outcome questions (response variables) to causative questions (explanatory variables).

4.1. Assumed Constraints

This study assumes that the learning outcomes at a university (Question Group 37) are influenced by opinions about university education (Question Group 32), attitudes towards learning activities (Question Group 33), and class experiences (Question Group 35). This is shown in Fig.1. From this assumption, we apply a constraint where directed arcs from the explanatory variables to the response variables are permitted but the reverse is prohibited. On the other hand, the following constraints about directed arcs among the explanatory variables can be considered:

- **constraint 1**: arcs among the explanatory variables belonging to the same group or different groups are permitted.

Table 1. Average and 95% confidence interval of the log-loss function values for each pair of constraints and score metrics.

		dist.	constraint 1			constraint 2			constraint 3			constraint 4		
			BIC	BDe	BDs	BIC	BDe	BDs	BIC	BDe	BDs	BIC	BDe	BDs
X_{32}	avg.	0.61	0.60	0.60	0.60	0.60	0.60	0.60	0.61	0.61	0.61	0.61	0.61	0.61
	95% C. I.	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
X_{33}	avg.	1.19	1.19	1.18	1.18	1.15	1.15	1.15	1.19	1.18	1.18	1.17	1.17	1.17
	95% C. I.	0.04	0.08	0.07	0.07	0.05	0.05	0.05	0.08	0.07	0.07	0.05	0.05	0.05
X_{35}	avg.	1.25	1.19	1.20	1.20	1.20	1.21	1.21	1.19	1.20	1.20	1.24	1.24	1.24
	95% C. I.	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03

- **constraint 2**: arcs among explanatory variables belonging to different groups are permitted, but arcs among explanatory variables belonging to the same group are prohibited.
- **constraint 3**: arcs among explanatory variables belonging to different groups are prohibited, but arcs among explanatory variables belonging to the same group are permitted.
- **constraint 4**: arcs among explanatory variables belonging to the same group or different groups are prohibited.

4.2. Implementation

In this study, bnlearn and gRain, R packages for Bayesian network learning and inference, are used in order to obtain the DAG from the questionnaire (observed data) and to calculate any conditional probabilities. The R, bnlearn, and gRain versions used are 3.4.3, 4.3, and 1.3.0, respectively. The bnlearn package provides both score-based and constraint-based learning algorithms. There is a case where the constraint-based algorithms return a partially directed graph, but the hill-climbing, a score-based learning algorithm, returns a completely directed graph. This study, thus, applies hill-climbing, and the parameters used as follows: the *restart* parameter, the number of random restarts, is 10,000 and the *perturb* parameter, the number of attempts to randomly insert/remove/invert an arc on every random restart, is 10. The candidate applied score metrics are Bayesian information Criterion (BIC)[18], Bayesian Dirichlet equivalence (BDe)[19], and Bayesian Dirichlet sparse (BDs)[20].

Calculation of the $P(X|Y=e)$, called *inference*, is performed by using the gRain package. The e is evidence, one of the possible outcomes for Y. For the inference of the Bayesian network, there are two approaches: strict inference and approximate inference. The former outputs the same results for each inference, while the latter outputs different results for each inference but the computational cost is lower than for the former. This study uses the strict approach.

4.3. Selection of Constraints and Score Metric

Suitable constraints and a score metric for the network learning are determined in terms of the prediction accuracy. This study concerns $P(X|Y)$ where $X \in \{X_{32} \cup X_{33} \cup X_{35}\}$, $X_{32} = \{X_{32.1}, \cdots, X_{32.10}\}$, $X_{33} = \{X_{33.1}, \cdots, X_{33.26}\}$, $X_{35} = \{X_{35.1}, \cdots, X_{35.19}\}$, and $Y \in \{X_{37.1}, \cdots, X_{37.28}\}$. The X_i corresponds to question i.

Prediction accuracy is measured using 10-fold cross validation, and the applied measure is the log-loss function s. Its definition is

$$s(X,Y) = -\frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{n} b_{ij}\log P(X=r_j|Y), \quad . \quad . \quad (5)$$

where m is the number of training examples, n is the number of possible outcomes for X, r_j is the j-th possible outcome for X, and b_{ij} is one if r_j equals to the outcome of X for the i-th training example, otherwise it is zero. The evidence of Y is also given by the i-th training example. From the definition, if $s(X,Y)$ differs by 0.01, it means that the predicted probability $P(X=r|Y)$ for the ground truth r is $e^{0.01} \approx 1.01$ times higher.

The Table 1 shows the averages and 95% confidence intervals of the log-loss function outputs for each pair of constraints and score metrics. The values are calculated for each question group, i.e., $s(Z_{32},Y)$ for all $Z_{32} \in X_{32}$, $s(Z_{33},Y)$ for all $Z_{33} \in X_{33}$, and $s(Z_{35},Y)$ for all $Z_{35} \in X_{35}$. The values in the column of "dist." are determined by: calculating a probability for each variable based on counting the training examples; apply the log-loss function to test examples. According to Table 1, the pair constraint 2 and BIC achieves the best prediction performance, because it has the best averages for Q_{32} and Q_{33}, second best average for Q_{35}, and best confidence intervals for Q_{32}, Q_{33}, and Q_{35}. On the other hand, the numbers of possible responses for Question Groups 32, 33, and 35 are 2, 4, and 4, respectively. If all responses are obtained randomly, then the function s outputs $-\log\frac{1}{2} = 0.69$ or $-\log\frac{1}{4} = 1.39$. From the random and "dist." cases, the prediction accuracy calculated with constraint 2 and BIC is reasonable.

5. Analysis with DAG and CPT

The subset of variables within Question Group 37 can be categorized into six abilities. In terms of these abilities, this section identifies and discusses factors influencing learning outcomes at a university based on a learned DAG and CPTs. The six abilities are as follows:

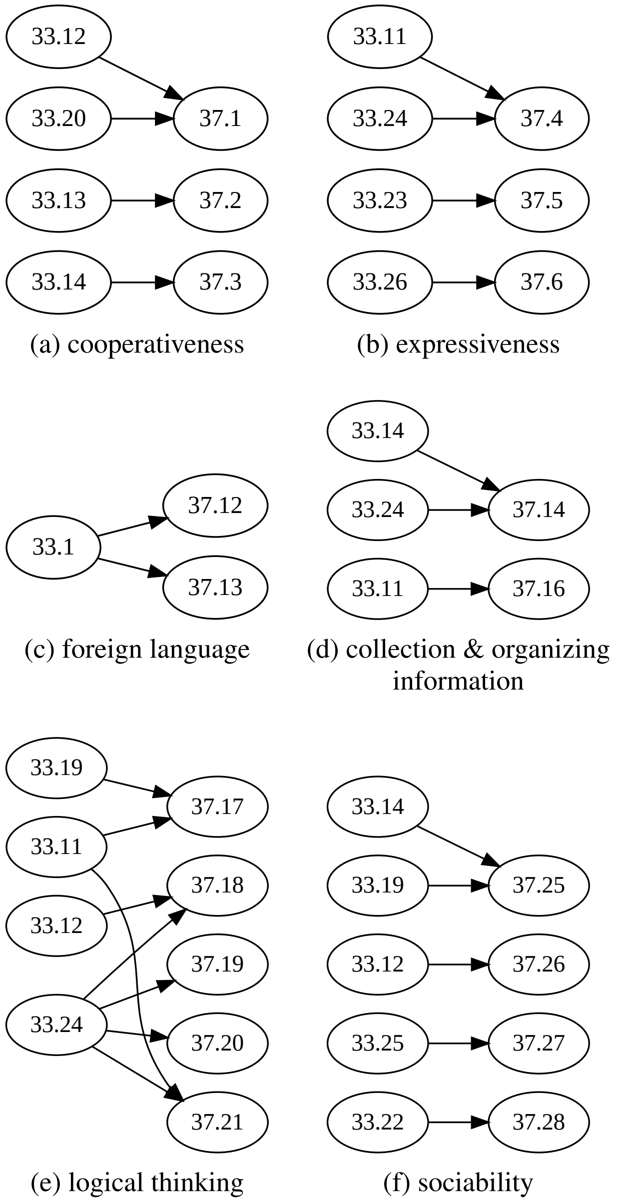

Fig. 2. Learned network for questions related to the six abilities within Question Group 33.

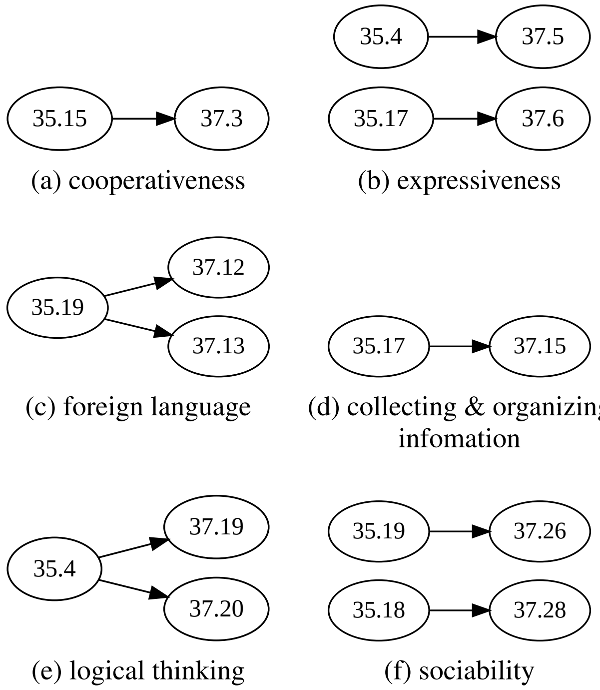

Fig. 3. Learned network for questions related to the six abilities within Question Group 35.

(a) **cooperativeness**: "ability to cooperate with others" ($X_{37.1}$); "ability to take initiative and act, and keep a group together" ($X_{37.2}$); "ability to arrange your own thoughts based on different opinions and positions" ($X_{37.3}$).

(b) **expressiveness**: "ability to write your own knowledge and ideas logically in sentences" ($X_{37.4}$); "ability to express your own knowledge and ideas with figures and numbers" ($X_{37.5}$); "ability to make documents and presentations using computers" ($X_{37.6}$).

(c) **foreign language**: "ability to read and write in foreign languages" ($X_{37.12}$); "ability to listen and converse in foreign languages" ($X_{37.13}$).

(d) **collecting and organizing information**: "ability to understand information in the literature and materials accurately" ($X_{37.14}$); "ability to create, organize, and analyze data using computers" ($X_{37.15}$); "ability to sort out diverse information accurately" ($X_{37.16}$).

(e) **logical thinking**: "ability to think critically and multilaterally" ($X_{37.17}$); "ability to analyze the current situation and discover problems and issues" ($X_{37.18}$); "ability to use mathematical expressions, figures, and graphs in order to solve problems" ($X_{37.19}$); "ability to plan and implement experiments and surveys appropriately for hypothesis verification and information gathering" ($X_{37.20}$); "ability to solve problems logically" ($X_{37.21}$).

(f) **sociability**: "ability to understand and respect the diversity of societies and cultures" ($X_{37.25}$); "an international perspective" ($X_{37.26}$); "ability to act according to norms and rules of society" ($X_{37.27}$); "ability to participate actively in social activities" ($X_{37.28}$).

5.1. Learned Networks for the Six Abilities

The network learning is performed with all responses of the fact-finding survey on university student life and learning. The constraint and score metric used are constraint 2 and BIC. The learned DAG for the six abilities is divided and shown in Fig.2 and Fig.3. There are no arcs from the variables within Question Group 32 to variables within Question Group 37; hence, the causal relationships among these variables can not be confirmed. On the other hand, there are arcs from the variables within Question Groups 33 and 35 to variables within Question Group 37. From Fig.2 and Fig.3, the factors, which influencing each ability, are identified as follows:

(a) **cooperativeness**: "contribute actively in group work and discussions" ($X_{33.12}$); "engage the facilitator voluntarily in group work and discussions" ($X_{33.13}$); "give consideration to different opinions and positions in group work and discussions" ($X_{33.14}$); "study with friends outside of group work" ($X_{33.20}$); "classes that give opportunities for discussions" ($X_{35.15}$).

(b) **expressiveness**: "express your own opinions in group work and discussions" ($X_{33.11}$); "make a study plan" ($X_{33.23}$); "willingly study continuously" ($X_{33.24}$); "actively work on your graduation thesis and research" ($X_{33.26}$); "classes that give opportunities for experimentation or research" ($X_{35.4}$); "classes that use computers and the Internet" ($X_{35.17}$).

(c) **foreign language**: "study in advance for classes" ($X_{33.1}$); "non-language classes done in foreign languages" ($X_{35.19}$).

(d) **collecting and organizing information**: "express your own opinions in group work and discussions" ($X_{33.11}$); "give consideration to different opinions and positions in group work and discussions" ($X_{33.14}$); "willingly study continuously" ($X_{33.24}$); "classes that use computers and the Internet" ($X_{35.17}$).

(e) **logical thinking**: "express your own opinions in group work and discussions" ($X_{33.11}$); "contribute actively in group work and discussions" ($X_{33.12}$); "voluntarily study things you are interested in regardless of classes" ($X_{33.19}$); "willingly study continuously" ($X_{33.24}$); "classes that give opportunities for experimentation or research" ($X_{35.4}$).

(f) **sociability**: "contribute actively in group work and discussions" ($X_{33.12}$); "give consideration to different opinions and positions in group work and discussions" ($X_{33.14}$); "voluntarily study things you are interested in regardless of classes" ($X_{33.19}$); "attend schools other than the university" ($X_{33.22}$); "try to get the best grade possible" ($X_{33.25}$); "classes to think about your career and aptitude" ($X_{35.18}$); "non-language classes done in foreign languages" ($X_{35.19}$).

Table 2. CPTs for cooperativeness.

$P(X_{33.12}, X_{33.20}|X_{37.1})$

			$X_{33.20}$			
			1	2	3	4
$X_{37.1}=1$	$X_{33.12}$	1	**0.28**	**0.11**	**0.06**	0.02
		2	**0.18**	**0.07**	0.04	0.03
		3	**0.07**	0.05	0.02	0.00
		4	0.03	0.01	0.02	0.00
$X_{37.1}=2$	$X_{33.12}$	1	0.06	**0.10**	**0.08**	0.01
		2	**0.15**	**0.23**	**0.12**	0.02
		3	**0.07**	**0.07**	0.06	0.01
		4	0.02	0.01	0.01	0.00
$X_{37.1}=3$	$X_{33.12}$	1	0.02	0.05	0.04	0.01
		2	0.06	**0.15**	**0.18**	0.03
		3	0.06	**0.14**	**0.15**	0.03
		4	0.01	0.03	0.03	0.01
$X_{37.1}=4$	$X_{33.12}$	1	0.01	0.02	0.04	0.02
		2	0.05	0.06	**0.10**	**0.07**
		3	0.05	**0.11**	**0.16**	**0.08**
		4	0.03	**0.07**	**0.08**	0.05

$P(X_{33.13}|X_{37.2})$

	$X_{33.13}$			
	1	2	3	4
$X_{37.2}=1$	**0.61**	**0.35**	0.03	0.01
$X_{37.2}=2$	0.18	**0.68**	0.12	0.01
$X_{37.2}=3$	0.09	**0.41**	**0.43**	0.08
$X_{37.2}=4$	0.06	0.19	**0.35**	**0.40**

$P(X_{33.14}, X_{35.15}|X_{37.3})$

			$X_{35.15}$			
			1	2	3	4
$X_{37.3}=1$	$X_{33.14}$	1	**0.42**	**0.08**	0.06	0.01
		2	**0.13**	0.05	**0.08**	0.00
		3	0.05	0.02	0.05	0.01
		4	0.02	0.01	0.02	0.00
$X_{37.3}=2$	$X_{33.14}$	1	**0.09**	0.06	0.04	0.00
		2	**0.12**	**0.21**	**0.11**	0.02
		3	0.06	**0.13**	**0.11**	0.02
		4	0.01	0.01	0.01	0.00
$X_{37.3}=3$	$X_{33.14}$	1	0.04	0.02	0.02	0.00
		2	0.05	**0.11**	**0.11**	0.02
		3	**0.07**	**0.15**	**0.26**	0.06
		4	0.01	0.02	0.04	0.02
$X_{37.3}=4$	$X_{33.14}$	1	0.04	0.00	0.02	0.00
		2	0.03	0.03	0.05	0.02
		3	0.03	**0.10**	**0.21**	**0.14**
		4	0.02	**0.06**	**0.11**	**0.13**

5.2. Discussion with CPTs

The CPTs for the six abilities are shown from Table 2 to Table 7. These tables only treat the variables in Fig.2 and Fig.3, since the conditional variables do not affect the result variables in other cases, i.e., $P(X,Y|Z) = P(X,Y)$ where the pairs (X, Z) and (Y, Z) are conditionally independent. The numbers of possible outcomes for the variables within Question Groups 33, 35, and 37 are all four; thus, for $X,Y \in \{X_{33} \cup X_{35}\}$, $Z \in X_{37}$, possible outcomes r and r', and evidence e, $P(X=r|Z=e)$ and $P(X=r,Y=r'|Z=e)$ are equal to $\frac{1}{4}=0.25$ and $\frac{1}{16}=0.0625$, respectively, if the distributions of $P(X|Z=e)$ and $P(X,Y|Z=e)$ are uniform. From Table 2 to Table 7, the probabilities that satisfy the following conditions are either only bolded or bolded and underlined: bolded only for when a probability is greater than the uniform case, $P(X=r|Z=e) > 0.25$ or $P(X=r,Y=r'|Z=e) > 0.0625$; bolded and underlined for when a probability is 1.5 times greater than the uniform case, $P(X=r|Z=e) > 0.25 \times 1.5$ or $P(X=r,Y=r'|Z=e) > 0.0625 \times 1.5$.

Table 3. CPTs for expressiveness.

$P(X_{33.11}, X_{33.24}|X_{37.4})$

			$X_{33.24}$			
			1	2	3	4
$X_{37.4} = 1$	$X_{33.11}$	1	**0.20**	**0.12**	**0.10**	0.04
		2	**0.08**	**0.10**	**0.10**	0.01
		3	**0.07**	0.05	0.04	0.02
		4	0.02	0.01	0.02	0.01
$X_{37.4} = 2$	$X_{33.11}$	1	0.03	**0.09**	**0.07**	0.02
		2	**0.06**	**0.20**	**0.16**	0.03
		3	0.05	**0.11**	**0.11**	0.02
		4	0.00	0.03	0.02	0.00
$X_{37.4} = 3$	$X_{33.11}$	1	0.01	0.04	0.06	0.01
		2	0.04	**0.11**	**0.18**	0.04
		3	0.03	**0.13**	**0.21**	0.04
		4	0.01	0.03	0.05	0.01
$X_{37.4} = 4$	$X_{33.11}$	1	0.02	0.01	0.05	0.02
		2	0.02	0.04	**0.10**	**0.06**
		3	0.02	**0.11**	**0.18**	**0.12**
		4	0.02	0.04	**0.12**	**0.07**

$P(X_{33.23}, X_{35.4}|X_{37.5})$

			$X_{35.4}$			
			1	2	3	4
$X_{37.5} = 1$	$X_{33.23}$	1	**0.20**	**0.07**	**0.10**	0.03
		2	**0.19**	**0.07**	0.06	0.03
		3	**0.13**	0.03	0.03	0.01
		4	0.03	0.01	0.00	0.00
$X_{37.5} = 2$	$X_{33.23}$	1	0.05	0.06	0.05	0.02
		2	**0.14**	**0.13**	**0.13**	0.04
		3	**0.11**	**0.09**	**0.07**	0.03
		4	0.03	0.02	0.02	0.01
$X_{37.5} = 3$	$X_{33.23}$	1	0.02	0.03	0.05	0.03
		2	**0.08**	**0.08**	**0.15**	**0.07**
		3	**0.09**	**0.09**	**0.16**	**0.06**
		4	0.02	0.02	0.04	0.01
$X_{37.5} = 4$	$X_{33.23}$	1	0.02	0.03	0.04	0.06
		2	0.06	0.02	0.06	**0.11**
		3	0.06	**0.07**	**0.12**	**0.12**
		4	0.04	0.05	**0.08**	**0.07**

$P(X_{33.26}, X_{35.17}|X_{37.6})$

			$X_{35.17}$			
			1	2	3	4
$X_{37.6} = 1$	$X_{33.26}$	1	**0.18**	**0.17**	**0.20**	0.06
		2	**0.08**	0.05	0.05	0.02
		3	0.04	0.04	0.03	0.03
		4	0.04	0.01	0.02	0.01
$X_{37.6} = 2$	$X_{33.26}$	1	0.03	**0.08**	**0.13**	0.04
		2	0.04	**0.11**	**0.14**	0.04
		3	0.05	**0.10**	**0.14**	0.03
		4	0.01	0.02	0.04	0.01
$X_{37.6} = 3$	$X_{33.26}$	1	0.01	0.05	**0.11**	0.05
		2	0.02	0.05	**0.14**	0.06
		3	0.03	**0.09**	**0.20**	**0.08**
		4	0.01	0.03	0.06	0.02
$X_{37.6} = 4$	$X_{33.26}$	1	0.01	0.02	**0.08**	**0.10**
		2	0.01	0.02	**0.08**	**0.10**
		3	0.02	0.04	**0.13**	**0.15**
		4	0.02	0.03	**0.10**	**0.09**

Table 4. CPTs for foreign language.

$P(X_{33.1}, X_{35.19}|X_{37.12})$

			$X_{35.19}$			
			1	2	3	4
$X_{37.12} = 1$	$X_{33.1}$	1	**0.28**	**0.08**	**0.07**	0.03
		2	**0.23**	0.06	0.04	0.01
		3	**0.15**	0.02	0.01	0.00
		4	0.02	0.01	0.00	0.00
$X_{37.12} = 2$	$X_{33.1}$	1	**0.11**	**0.07**	0.05	0.01
		2	**0.23**	**0.14**	**0.08**	0.02
		3	**0.14**	**0.08**	0.04	0.01
		4	0.01	0.01	0.01	0.00
$X_{37.12} = 3$	$X_{33.1}$	1	**0.08**	0.05	0.05	0.02
		2	**0.16**	**0.08**	**0.12**	0.04
		3	**0.13**	**0.08**	**0.11**	0.03
		4	0.03	0.01	0.01	0.00
$X_{37.12} = 4$	$X_{33.1}$	1	0.05	0.02	0.03	**0.06**
		2	**0.07**	0.03	0.06	**0.12**
		3	**0.13**	**0.09**	**0.07**	**0.13**
		4	0.04	0.04	0.03	0.03

$P(X_{33.1}, X_{35.19}|X_{37.13})$

			$X_{35.19}$			
			1	2	3	4
$X_{37.13} = 1$	$X_{33.1}$	1	**0.27**	**0.09**	**0.06**	0.02
		2	**0.24**	0.06	0.04	0.01
		3	**0.14**	0.02	0.01	0.00
		4	0.02	0.00	0.00	0.00
$X_{37.13} = 2$	$X_{33.1}$	1	**0.10**	0.06	0.05	0.02
		2	**0.22**	**0.14**	**0.09**	0.02
		3	**0.14**	**0.08**	0.04	0.01
		4	0.01	0.01	0.00	0.00
$X_{37.13} = 3$	$X_{33.1}$	1	**0.07**	0.05	0.05	0.02
		2	**0.15**	**0.08**	**0.11**	0.04
		3	**0.13**	**0.08**	**0.12**	0.03
		4	0.03	0.01	0.01	0.00
$X_{37.13} = 4$	$X_{33.1}$	1	0.04	0.03	0.05	0.06
		2	0.05	0.02	**0.07**	**0.14**
		3	**0.12**	**0.11**	0.06	**0.12**
		4	0.03	0.04	0.03	0.03

5.2.1. Cooperativeness

The $P(X_{33.12}, X_{33.20}|X_{37.1})$, $P(X_{33.13}|X_{37.2})$, and $P(X_{33.14}, X_{35.15}|X_{37.3})$ distributions shift to high agreement/experience levels in Table 2, as self-knowledge levels about the acquired abilities increase:

- both "contribute actively in group work and discussions" ($X_{33.12}$) and "study with friends outside of group work" ($X_{33.20}$) correlate "ability to cooperate with others" ($X_{37.1}$) positively.

- "engage the facilitator voluntarily in group work and discussions" ($X_{33.13}$) correlates "ability to take initiative and act, and keep a group together" ($X_{37.2}$) positively.

- both "give consideration to different opinions and positions in group work and discussions" ($X_{33.14}$)

Table 5. CPTs for collection and organizing information.

$P(X_{33.14}, X_{33.24} | X_{37.14})$

			$X_{33.24}$			
			1	2	3	4
$X_{37.14} = 1$	$X_{33.14}$	1	**0.15**	**0.08**	**0.06**	0.05
		2	**0.12**	**0.11**	**0.09**	0.03
		3	**0.10**	**0.08**	0.05	0.02
		4	0.03	0.02	0.01	0.01
$X_{37.14} = 2$	$X_{33.14}$	1	0.03	**0.10**	**0.07**	0.02
		2	**0.07**	**0.18**	**0.13**	0.02
		3	0.06	**0.15**	**0.12**	0.03
		4	0.01	0.02	0.01	0.00
$X_{37.14} = 3$	$X_{33.14}$	1	0.01	0.03	**0.06**	0.01
		2	0.03	**0.08**	**0.16**	0.04
		3	0.05	**0.15**	**0.23**	0.05
		4	0.01	0.03	0.04	0.01
$X_{37.14} = 4$	$X_{33.14}$	1	0.01	0.01	0.04	0.02
		2	0.01	0.03	**0.07**	**0.07**
		3	0.04	**0.10**	**0.20**	**0.14**
		4	0.03	0.06	**0.10**	0.05

$P(X_{35.17} | X_{37.15})$

	$X_{35.17}$			
	1	2	3	4
$X_{37.15} = 1$	**0.38**	0.24	**0.32**	0.06
$X_{37.15} = 2$	0.10	**0.34**	**0.44**	0.12
$X_{37.15} = 3$	0.06	0.19	**0.53**	0.22
$X_{37.15} = 4$	0.04	0.10	**0.35**	**0.51**

$P(X_{33.11} | X_{37.16})$

	$X_{33.11}$			
	1	2	3	4
$X_{37.16} = 1$	**0.54**	0.25	0.14	0.08
$X_{37.16} = 2$	0.22	**0.48**	**0.26**	0.04
$X_{37.16} = 3$	0.11	**0.38**	**0.42**	0.09
$X_{37.16} = 4$	0.08	0.23	**0.40**	**0.29**

and "classes that give opportunities for discussions" ($X_{35.15}$) correlate "ability to arrange your own thoughts based on different opinions and positions" ($X_{37.3}$) positively.

5.2.2. Expressiveness

The $P(X_{33.11}, X_{33.24} | X_{37.4})$, $P(X_{33.23}, X_{35.4} | X_{37.5})$, and $P(X_{33.26}, X_{35.17} | X_{37.6})$ distributions shift to high agreement/experience levels in Table 3, as self-knowledge levels about the acquired abilities increase. If $X_{37.6} \geq 3$, the $P(X_{33.26}, X_{35.17} | X_{37.6})$ distribution is biased to $X_{35.17} \geq 3$:

- both "express your own opinions in group work and discussions" ($X_{33.11}$) and "willingly study continuously" ($X_{33.24}$) correlate "ability to write your own knowledge and ideas logically in sentences" ($X_{37.4}$) positively.
- both "make a study plan" ($X_{33.23}$) and "classes that give opportunities for experimentation or research" ($X_{35.4}$) correlate "ability to express your own knowledge and ideas with figures and numbers" ($X_{37.5}$) positively.
- both "actively work on your graduation thesis and research" ($X_{33.26}$) and "classes that use computers and the Internet" ($X_{35.17}$) correlate "ability to make documents and presentations using computers" ($X_{37.6}$) positively, and the effect of the latter factor is especially high.

5.2.3. Foreign Language

The $P(X_{33.1}, X_{35.19} | X_{37.12})$ and $P(X_{33.1}, X_{35.19} | X_{37.13})$ distributions shift to high agreement/experience levels in Table 4, as self-knowledge levels about the acquired abilities increase. The degrees of those shifts are gradual, and the probabilities for the $X_{33.1} = 4$ case are less. These suggests that: both "study in advance for classes" ($X_{33.1}$) and "non-language classes done in foreign languages" ($X_{35.19}$) correlate both "ability to read and write in foreign languages" ($X_{37.12}$) and "ability to listen and converse in foreign languages" ($X_{37.13}$) positively, but the effects of the former factor are relatively weak.

5.2.4. Collecting and Organizing Information

The $P(X_{33.14}, X_{33.24} | X_{37.14})$, $P(X_{35.17} | X_{37.15})$, and $P(X_{33.11} | X_{37.16})$ distributions shift to high agreement/experience levels in Table 5, as self-knowledge levels about the acquired abilities increase:

- both "give consideration to different opinions and positions in group work and discussions" ($X_{33.14}$) and "willingly study continuously" ($X_{33.24}$) correlate "ability to understand information in the literature and materials accurately" ($X_{37.14}$) positively.
- "classes that use computers and the Internet" ($X_{35.17}$) correlates "ability to create, organize, and analyze data using computers" ($X_{37.15}$) positively.
- "express your own opinions in group work and discussions" ($X_{33.11}$) correlates "ability to sort out diverse information accurately" ($X_{37.16}$) positively.

5.2.5. Logical Thinking

The $P(X_{33.11}, X_{33.19} | X_{37.17})$, $P(X_{33.12}, X_{33.24} | X_{37.18})$, $P(X_{33.24}, X_{35.4} | X_{37.19})$, $P(X_{33.24}, X_{35.4} | X_{37.20})$, and $P(X_{33.11}, X_{33.24} | X_{37.21})$ distributions shift to high agreement/experience levels in Table 6, as self-knowledge levels about the acquired abilities increase:

- both "express your own opinions in group work and discussions" ($X_{33.11}$) and "voluntarily study things you are interested in regardless of classes" ($X_{33.19}$) correlate "ability to think critically and multilaterally" ($X_{37.17}$) positively.

Table 6. CPTs for logical thinking.

$P(X_{33.11}, X_{33.19} | X_{37.17})$

			$X_{33.19}$			
			1	2	3	4
$X_{37.17} = 1$	$X_{33.11}$	1	**0.14**	**0.10**	**0.15**	**0.06**
		2	**0.09**	**0.07**	**0.16**	0.02
		3	0.03	0.06	0.05	0.02
		4	0.01	0.02	0.02	0.00
$X_{37.17} = 2$	$X_{33.11}$	1	0.03	**0.08**	**0.08**	0.02
		2	0.06	**0.19**	**0.17**	0.05
		3	0.04	**0.09**	**0.11**	0.03
		4	0.00	0.02	0.03	0.01
$X_{37.17} = 3$	$X_{33.11}$	1	0.01	0.04	0.05	0.02
		2	0.03	**0.10**	**0.17**	**0.07**
		3	0.03	**0.11**	**0.20**	**0.08**
		4	0.01	0.03	0.04	0.02
$X_{37.17} = 4$	$X_{33.11}$	1	0.01	0.01	0.05	0.03
		2	0.01	0.04	**0.11**	**0.11**
		3	0.02	**0.09**	**0.16**	**0.15**
		4	0.02	0.04	**0.08**	**0.09**

$P(X_{33.12}, X_{33.24} | X_{37.18})$

			$X_{33.24}$			
			1	2	3	4
$X_{37.18} = 1$	$X_{33.12}$	1	**0.14**	**0.11**	**0.11**	0.05
		2	**0.14**	**0.10**	**0.08**	0.03
		3	**0.06**	**0.07**	0.05	0.02
		4	0.02	0.02	0.00	0.01
$X_{37.18} = 2$	$X_{33.12}$	1	0.02	**0.10**	**0.09**	0.02
		2	**0.07**	**0.21**	**0.18**	0.03
		3	0.06	**0.08**	**0.08**	0.01
		4	0.01	0.02	0.02	0.00
$X_{37.18} = 3$	$X_{33.12}$	1	0.01	0.03	**0.07**	0.02
		2	0.03	**0.11**	**0.21**	0.05
		3	0.03	**0.12**	**0.19**	0.04
		4	0.01	0.03	0.05	0.01
$X_{37.18} = 4$	$X_{33.12}$	1	0.01	0.02	0.05	0.02
		2	0.03	0.04	**0.11**	**0.10**
		3	0.05	**0.09**	**0.16**	**0.11**
		4	0.03	0.06	**0.09**	0.05

$P(X_{33.24}, X_{35.4} | X_{37.19})$

			$X_{35.4}$			
			1	2	3	4
$X_{37.19} = 1$	$X_{33.24}$	1	**0.14**	0.05	**0.08**	0.02
		2	**0.18**	0.06	0.05	0.02
		3	**0.21**	0.06	0.05	0.01
		4	0.06	0.02	0.01	0.00
$X_{37.19} = 2$	$X_{33.24}$	1	0.03	0.04	0.03	0.01
		2	**0.11**	**0.12**	**0.11**	0.03
		3	**0.13**	**0.14**	**0.11**	0.04
		4	0.03	0.02	0.03	0.01
$X_{37.19} = 3$	$X_{33.24}$	1	0.01	0.02	0.03	0.02
		2	**0.07**	0.06	**0.12**	0.05
		3	**0.11**	**0.11**	**0.20**	**0.08**
		4	0.02	0.03	0.05	0.02
$X_{37.19} = 4$	$X_{33.24}$	1	0.01	0.02	0.03	0.03
		2	0.03	0.03	0.05	**0.10**
		3	**0.08**	0.06	**0.13**	**0.17**
		4	0.05	0.06	**0.07**	**0.09**

$P(X_{33.24}, X_{35.4} | X_{37.20})$

			$X_{35.4}$			
			1	2	3	4
$X_{37.20} = 1$	$X_{33.24}$	1	**0.15**	0.06	**0.07**	0.01
		2	**0.20**	0.05	0.04	0.02
		3	**0.22**	0.05	0.03	0.01
		4	**0.07**	0.01	0.02	0.01
$X_{37.20} = 2$	$X_{33.24}$	1	0.03	0.04	0.03	0.01
		2	**0.13**	**0.12**	**0.11**	0.03
		3	**0.14**	**0.14**	**0.09**	0.02
		4	0.03	0.03	0.02	0.00
$X_{37.20} = 3$	$X_{33.24}$	1	0.01	0.01	0.03	0.02
		2	0.05	0.06	**0.12**	0.06
		3	**0.11**	**0.11**	**0.22**	**0.09**
		4	0.03	0.03	0.05	0.02
$X_{37.20} = 4$	$X_{33.24}$	1	0.01	0.01	0.03	0.04
		2	0.01	0.02	0.05	**0.09**
		3	0.05	0.04	**0.13**	**0.20**
		4	0.05	**0.07**	**0.09**	**0.10**

$P(X_{33.11}, X_{33.24} | X_{37.21})$

			$X_{33.24}$			
			1	2	3	4
$X_{37.21} = 1$	$X_{33.11}$	1	**0.19**	**0.09**	**0.09**	0.04
		2	**0.13**	**0.10**	**0.08**	0.03
		3	**0.07**	**0.07**	0.04	0.02
		4	0.00	0.02	0.03	0.01
$X_{37.21} = 2$	$X_{33.11}$	1	0.03	**0.10**	**0.07**	0.01
		2	**0.07**	**0.19**	**0.15**	0.03
		3	0.04	**0.13**	**0.11**	0.02
		4	0.01	0.02	0.01	0.00
$X_{37.21} = 3$	$X_{33.11}$	1	0.01	0.03	0.05	0.01
		2	0.03	**0.11**	**0.19**	0.04
		3	0.04	**0.11**	**0.22**	0.05
		4	0.01	0.03	0.05	0.01
$X_{37.21} = 4$	$X_{33.11}$	1	0.01	0.02	**0.06**	0.02
		2	0.02	0.03	**0.10**	0.06
		3	0.02	**0.12**	**0.16**	**0.11**
		4	0.02	0.04	**0.13**	**0.09**

Table 7. CPTs for sociability.

$P(X_{33.14}, X_{33.19}|X_{37.25})$

			$X_{33.19}$ 1	2	3	4
$X_{37.25}=1$	$X_{33.14}$	1	**0.07**	$\underline{\mathbf{0.10}}$	$\underline{\mathbf{0.12}}$	**0.09**
		2	0.06	**0.07**	$\underline{\mathbf{0.11}}$	0.05
		3	0.06	0.05	**0.06**	0.04
		4	0.03	0.03	0.03	0.01
$X_{37.25}=2$	$X_{33.14}$	1	0.02	**0.07**	**0.08**	0.03
		2	0.03	$\underline{\mathbf{0.15}}$	$\underline{\mathbf{0.15}}$	0.05
		3	0.06	$\underline{\mathbf{0.13}}$	$\underline{\mathbf{0.14}}$	0.04
		4	0.00	0.03	0.02	0.01
$X_{37.25}=3$	$X_{33.14}$	1	0.01	0.03	0.05	0.01
		2	0.03	**0.08**	$\underline{\mathbf{0.15}}$	0.06
		3	0.03	$\underline{\mathbf{0.15}}$	$\underline{\mathbf{0.24}}$	$\underline{\mathbf{0.10}}$
		4	0.00	0.02	0.04	0.01
$X_{37.25}=4$	$X_{33.14}$	1	0.00	0.01	0.03	0.04
		2	0.01	0.02	**0.08**	$\underline{\mathbf{0.10}}$
		3	0.02	**0.09**	$\underline{\mathbf{0.19}}$	$\underline{\mathbf{0.17}}$
		4	0.02	0.05	**0.09**	**0.07**

$P(X_{33.12}, X_{35.19}|X_{37.26})$

			$X_{35.19}$ 1	2	3	4
$X_{37.26}=1$	$X_{33.12}$	1	$\underline{\mathbf{0.22}}$	**0.06**	0.05	0.02
		2	$\underline{\mathbf{0.22}}$	0.04	0.04	0.01
		3	$\underline{\mathbf{0.18}}$	0.04	0.01	0.01
		4	0.06	0.01	0.01	0.00
$X_{37.26}=2$	$X_{33.12}$	1	**0.07**	0.05	0.03	0.02
		2	$\underline{\mathbf{0.26}}$	$\underline{\mathbf{0.13}}$	**0.08**	0.02
		3	$\underline{\mathbf{0.15}}$	**0.07**	0.04	0.01
		4	0.04	0.02	0.01	0.00
$X_{37.26}=3$	$X_{33.12}$	1	0.05	0.02	0.03	0.01
		2	$\underline{\mathbf{0.17}}$	$\underline{\mathbf{0.10}}$	$\underline{\mathbf{0.10}}$	0.04
		3	$\underline{\mathbf{0.15}}$	$\underline{\mathbf{0.10}}$	$\underline{\mathbf{0.11}}$	0.03
		4	0.04	0.02	0.03	0.01
$X_{37.26}=4$	$X_{33.12}$	1	0.04	0.03	0.02	0.02
		2	**0.07**	0.05	**0.07**	$\underline{\mathbf{0.09}}$
		3	$\underline{\mathbf{0.15}}$	**0.08**	**0.08**	$\underline{\mathbf{0.10}}$
		4	**0.07**	0.06	0.03	0.05

$P(X_{33.25}|X_{37.27})$

	$X_{33.25}$ 1	2	3	4
$X_{37.27}=1$	**0.33**	0.21	0.21	0.25
$X_{37.27}=2$	0.09	**0.32**	$\underline{\mathbf{0.42}}$	0.17
$X_{37.27}=3$	0.04	0.17	$\underline{\mathbf{0.50}}$	**0.29**
$X_{37.27}=4$	0.05	0.10	**0.33**	$\underline{\mathbf{0.51}}$

$P(X_{33.22}, X_{35.18}|X_{37.28})$

			$X_{35.18}$ 1	2	3	4
$X_{37.28}=1$	$X_{33.22}$	1	$\underline{\mathbf{0.35}}$	$\underline{\mathbf{0.22}}$	$\underline{\mathbf{0.14}}$	0.04
		2	**0.08**	0.04	0.03	0.01
		3	0.03	0.01	0.01	0.00
		4	0.03	0.01	0.00	0.00
$X_{37.28}=2$	$X_{33.22}$	1	$\underline{\mathbf{0.12}}$	$\underline{\mathbf{0.21}}$	$\underline{\mathbf{0.15}}$	0.02
		2	$\underline{\mathbf{0.09}}$	$\underline{\mathbf{0.13}}$	**0.08**	0.01
		3	0.05	0.05	0.02	0.01
		4	0.01	0.02	0.02	0.00
$X_{37.28}=3$	$X_{33.22}$	1	$\underline{\mathbf{0.10}}$	$\underline{\mathbf{0.12}}$	$\underline{\mathbf{0.16}}$	0.02
		2	**0.07**	**0.09**	$\underline{\mathbf{0.11}}$	0.03
		3	0.06	**0.07**	$\underline{\mathbf{0.10}}$	0.02
		4	0.01	0.02	0.03	0.00
$X_{37.28}=4$	$X_{33.22}$	1	$\underline{\mathbf{0.12}}$	$\underline{\mathbf{0.17}}$	$\underline{\mathbf{0.22}}$	$\underline{\mathbf{0.13}}$
		2	0.03	0.03	0.02	0.04
		3	0.01	0.03	0.05	0.02
		4	0.03	0.03	0.04	0.03

- both "contribute actively in group work and discussions" ($X_{33.12}$) and "willingly study continuously" ($X_{33.24}$) correlate "ability to analyze the current situation and discover problems and issues" ($X_{37.18}$) positively.

- both "willingly study continuously" ($X_{33.24}$) and "classes that give opportunities for experimentation or research" ($X_{35.4}$) correlate "ability to use mathematical expressions, figures, and graphs in order to solve problems" ($X_{37.19}$) positively.

- both "willingly study continuously" ($X_{33.24}$) and "classes that give opportunities for experimentation or research" ($X_{35.4}$) correlate "ability to plan and implement experiments and surveys appropriately for hypothesis verification and information gathering" ($X_{37.20}$) positively.

- both "express your own opinions in group work and discussions" ($X_{33.11}$) and "willingly study continuously" ($X_{33.24}$) correlate "ability to solve problems logically" ($X_{37.21}$) positively.

5.2.6. Sociability

The $P(X_{33.14}, X_{33.19}|X_{37.25})$ and $P(X_{33.25}|X_{37.27})$ distributions shift to high agreement levels in Table 7, as self-knowledge levels about the acquired abilities increase. The shift in the $P(X_{33.12}, X_{35.19}|X_{37.26})$ distribution is gradual. On the other hand, the $P(X_{33.22}, X_{35.18}|X_{37.28})$ distribution is biased towards $X_{33.22} = 1$ for all $X_{37.28}$, and the probabilities for the $X_{35.18} = 4$ case are smaller:

- both "give consideration to different opinions and positions in group work and discussions" ($X_{33.14}$) and "voluntarily study things you are interested in regardless of classes" ($X_{33.19}$) correlate "ability to understand and respect the diversity of societies and cultures” ($X_{37.25}$) positively.

- both “contribute actively in group work and discussions” ($X_{33.12}$) and “non-language classes done in foreign languages” ($X_{35.19}$) correlate “an international perspective” ($X_{37.26}$) positively.

- “try to get the best grade possible” ($X_{33.25}$) correlates “ability to act according to norms and rules of society” ($X_{37.27}$) positively.

- “attend schools other than the university” ($X_{33.22}$) correlates “ability to participate actively in social activities” ($X_{37.28}$) negatively, and “classes to think about your career and aptitude” ($X_{35.18}$) correlates “ability to participate actively in social activities” ($X_{37.28}$) positively, but effects of this factor are relatively weak.

6. Conclusion

This study performs a secondary analysis on a fact-finding survey on university student life and learning[2] with a Bayesian network. In this study, four question groups are used in the survey: (i) learning outcomes at the university; (ii) opinions about education in the university; (iii) attitude towards learning activities; and (iv) class experience. The questions within (i) are used as response variables and this subset is categorized according to six abilities: cooperativeness, expressiveness, foreign language, collecting and organizing information, logical thinking, and sociability. The questions within (ii), (iii), and (iv) are used as explanatory variables, and factors that influence the response variables are identified and analyzed.

Suitable constraints and a score metric for the network learning are determined in terms of the prediction accuracy. The learned DAG under the conditions suggests that two to seven explanatory variables influence each ability. In addition, there are no arcs from variables within the question group (ii) to variables within the question group (i); hence, causal relationships among these variables cannot be confirmed. In contrast, there are arcs from variables within the question groups (iii) and (iv) to variables within the question group (i). The effects of the identified influence factors are analyzed in terms of the conditional probabilities. The probability distributions of most factors shift to high agreement/experience levels, as self-knowledge levels about the acquired abilities increase, i.e., positive effects exist for most factors for each ability.

In future work, we will propose measures to improve the six-abilities based on the identified factors. We are especially interested in the kinds of learning environments that universities should provide, and we will take into account the results of this study.

Acknowledgements

The data for this secondary analysis, Fact-finding Survey on University Student Life and Learning, 2008, Benesse Educational Research and Development Institute (former name: Benesse Corporation), was provided by the Social Science Japan Data Archive, Center for Social Research and Data Archives, Institute of Social Science, The University of Tokyo.

References:

[1] Department of University Management and Policy Studies, The University of Tokyo, “Nationwide University Student Survey,” http://ump.p.u-tokyo.ac.jp/crump/cat77/cat82/post-6.html. [written in Japanese, 2017/07/13 accessed]
[2] Social Science Japan Data Archive, “Fact-finding Survey on University Student Life and Learning,” http://ssjda.iss.u-tokyo.ac.jp/Direct/gaiyo.php?lang=eng&eid=0721. [2017/07/13 accessed]
[3] National Institute for Educational Policy Research, “Survey on University Student Learning,” http://www.nier.go.jp/04_kenkyu_annai/pdf/gakushu-jittai_2014.pdf. [written in Japanese, 2017/07/13 accessed]
[4] Kanako Shima, Kazuma Oyamada, Shintaro Nakayama, and Masaya Yoshimasu, “Daigakukyoiku no Shitsuteki Jujitsuka – Jun-Specialist no Haishutsu ni Mukete –,” Policy Forum 2015, Inter-university Seminar for the Future of Japan, 2015. [written in Japanese]
[5] Hiroki Yasuda, “Gender Differences in Grades and Efforts to Class,” The Journal of Tokyo Keizai University: Economics, Vol.289, pp.85–101, 2016. [written in Japanese]
[6] Judea Pearl, “Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning,” Proceedings of 7th Annual Conference of the Cognitive Science Society, 1985.
[7] Gregory F. Cooper and Edward Herskovits, “A Bayesian Method for the Induction of Probabilistic Networks from Data,” Machine Learning, Vol. 9, No. 4, pp.309–347, 1992.
[8] Dimitris Margaritis, “Learning Bayesian Network Model Structure from Data,” Ph.D. thesis, School of Computer Science, Carnegie-Mellon University, 2003.
[9] David M. Chickering, “Learning Bayesian Networks is NP-Complete,” Learning from Data, Lecture Notes in Statistics, Vol.112, pp.121–130, 1996.
[10] Pathom Pumpuang, Anongnart Srivihok, Prasong Praneetpolgrang, and Somchai Numprasertchai, “Using Bayesian Network for Planning Course Registration Model for Undergraduate Students,” Proceedings of 2nd IEEE International Conference on Digital Ecosystems and Technologies, pp. 492–496, 2008.
[11] Ashkan Sharabiani, Fazle Karim, Anooshiravan Sharabiani, Mariya Atanasov, and Houshang Darabi, “An Enhanced Bayesian Network Model for Prediction of Students’ Academic Performance in Engineering Programs,” Proceedings of 2014 IEEE Global Engineering Education Conference, pp.832–837, 2014.
[12] Antonio Fernández, María Morales, Carmelo Rodríguez, and Antonio Salmerón, “A System for Relevance Analysis of Performance Indicators in Higher Education using Bayesian Networks,” Knowledge and Information Systems, Vol. 27, No. 3, pp.327–344, 2011.
[13] Eva Millán, Tomasz Loboda, and Jose Luis Pérez-de-la-Cruz, “Bayesian Networks for Student Model Engineering,” Computers & Education, Vol.55, No.4, pp.1663–1683, 2010.
[14] Ani Grubišić, Slavomir Stankov, and Ivan Peraić, “Ontology Based Approach to Bayesian Student Model Design,” Expert Systems with Applications, Vol. 40, No. 13, pp.5363–5371, 2013.
[15] Michalis Xenos, “Prediction and Assessment of Student Behaviour in Open and Distance Education in Computers using Bayesian Networks,” Computers & Education, Vol.43, No.4, pp.345–359, 2004.
[16] Patricio García, Analía Amandi, Silvia Schiaffino, and Marcelo Campo, “Evaluating Bayesian Networks’ Precision for Detecting Students’ Learning Styles,” Computers & Education, Vol. 49, No. 3, pp.794–808, 2007.
[17] Masashi Watanabe, Yasuo Tsuchida, Kota Kimura, and Michiko Tsubaki, “Analysis of the Educational Effectiveness Considering Individual Differences using Bayesian Network,” The European Conference on Educational Research 2009, 492, 2009.
[18] Gideon Schwarz, “Estimating the Dimension of a Model,” The Annals of Statistics, Vol. 6, No. 2, pp.461–464, 1978.
[19] David Heckerman, Dan Geiger, and David M. Chickering, “Learning Bayesian Networks: The Combination of Knowledge and Statistical Data,” Machine Learning, Vol. 20, No. 3, pp.197–243, 1995.
[20] Marco Scutari, “An Empirical-Bayes Score for Discrete Bayesian Networks,” Proceedings of 8th International Conference on Probabilistic Graphical Models, Vol. 52, pp.438–448, 2016.

Appendix A. Questionnaire Used

A.1. Opinions about Education in the University

A.1.1. Questions and Responses

32.1 Which course do you prefer?

1. A bird course, even if it is not very interesting.
2. An interesting course, even if it is difficult to get credits.

32.2 Which course do you prefer?

1. One where attendance and participation are emphasized in the grading.
2. One where examinations, papers, or reports are emphasized in the grading

32.3 Which course do you prefer?

1. One where the fundamentals are the main focus, even if there is less application and development content.
2. One where the application and development contents are the main focus, even if less fundamental.

32.4 Which opinion do you agree with?

1. There are many lecture-style courses where faculty members mainly lecture on knowledge and skills.
2. There are many practical-style courses where students study and report by themselves.

32.5 Which opinion do you agree with?

1. In university, it is better to acquire knowledge and skills in a wide range of fields.
2. In university, it is better to acquire knowledge and skills in specific, specialized fields.

32.6 Which opinion do you agree with?

1. Systematic learning is possible even if constraints on the choice of courses exist.
2. Free choice of courses is possible even if constraints on systematic learning exist.

32.7 Which opinion do you agree with?

1. In learning at the university, it is better that the students are coached by the faculty members via classes.
2. In learning at the university, it is better that the students learn by themselves.

32.8 Which opinion do you agree with?

1. It is better that the students decide on what they want to do in the future first and then take classes.
2. It is better that the students find out what they want to do in the future via the classes.

32.9 Which opinion do you agree with?

1. It is better that the faculty members interact actively with the students, even outside of classes.
2. It is better that the faculty members do not have to interact with the students more than necessary outside of classes.

32.10 Which opinion do you agree with?

1. It is better that the faculty members instruct the students on campus life.
2. It is better that the faculty members leave it to the students' autonomy with respect to campus life.

A.2. Attitude towards Learning Activities

A.2.1. Question: You

33.1 study in advance for classes.

33.2 always bring the textbooks, literature, and notes necessary for classes.

33.3 try not to be late for classes.

33.4 not give up registered courses halfway.

33.5 take notes on content not written on the blackboard during class.

33.6 not engage in private conversation during class.

33.7 ask faculty members about what you did not understand in class.

33.8 keep up with homework assigned in class.

33.9 review before submitting reports and tests.

33.10 ask questions and make comments actively in front of the whole class.

33.11 express your own opinions in group work and discussions.

33.12 contribute actively in group work and discussions.

33.13 engage the facilitator voluntarily in group work and discussions.

33.14 give consideration to different opinions and positions in group work and discussions.

33.15 review classes.

33.16 personally search for what you did not understand in class.

33.17 voluntarily study things that you interest you in class.

33.18 organize materials distributed in class.

33.19 voluntarily study things you are interested in regardless of classes.

33.20 study with friends outside of group work.

33.21 study in order to acquire qualifications and licenses.

33.22 attend schools other than the university.

33.23 make a study plan.

33.24 willingly study continuously.

33.25 try to get the best grade possible.

33.26 actively work on your graduation thesis and research.

A.2.2. Response Alternative

1. Strongly disagree
2. Weakly disagree
3. Weakly agree
4. Strongly agree

A.3. Class Experience

A.3.1. Question

35.1 Supplementary classes on subjects to study in high school.

35.2 Classes on how to study at a university.

35.3 Seminars with fewer students or practical-style classes.

35.4 Classes that give opportunities for experimentation or research.

35.5 Classes that have hands-on activities or practices outside the classroom.

35.6 Classes where students write comments or opinions about the content each time.
35.7 Classes where faculty members and students can communicate during class.
35.8 Classes where senior and junior students can communicate during class.
35.9 Classes where faculty members and students can communicate outside the classroom using the Internet and e-mail.
35.10 Classes that do tests other than at the end of the semester.
35.11 Classes that give assignments other than at the end of the semester.
35.12 Classes where submitted assignments are returned with comments by faculty members.
35.13 Classes where students' opinions and class evaluations are reflected.
35.14 Classes that do collaborative work such as group work.
35.15 Classes that give opportunities for discussions.
35.16 Classes that give opportunities for presentations.
35.17 Classes that use computers and the Internet.
35.18 Classes to think about your career and aptitude.
35.19 Non-language classes done in foreign languages.

A.3.2. Response Alternative

1. Rarely experienced
2. Not often experienced
3. Experienced to some extent
4. Experienced often

A.4. Learning Outcomes at the University

A.4.1. Question

37.1 Ability to cooperate with others.
37.2 Ability to take initiative and act, and keep a group together.
37.3 Ability to arrange your own thoughts based on different opinions and positions.
37.4 Ability to write your own knowledge and ideas logically in sentences.
37.5 Ability to express your own knowledge and ideas with figures and numbers.
37.6 Ability to make documents and presentations using computers.
37.7 Ability to actively try to acquire new knowledge and abilities.
37.8 Ability to set goals for yourself and act systematically.
37.9 Ability to control your own emotions well.
37.10 Ability to understand your own aptitude and abilities.
37.11 Ability to have confidence in yourself.
37.12 Ability to read and write in foreign languages.
37.13 Ability to listen and converse in foreign languages.
37.14 Ability to understand information in the literature and materials accurately.
37.15 Ability to create, organize, and analyze data using computers.
37.16 Ability to sort out diverse information accurately.
37.17 Ability to think critically and multilaterally.
37.18 Ability to analyze the current situation and discover problems and issues.
37.19 Ability to use mathematical expressions, figures, and graphs in order to solve problems.
37.20 Ability to plan and implement experiments and surveys appropriately for hypothesis verification and information gathering.
37.21 Ability to solve problems logically.
37.22 Ability to generate new ideas by thinking outside-the-box.
37.23 An education and general knowledge.
37.24 Basic knowledge and skills in a specialized field.
37.25 Ability to understand and respect the diversity of societies and cultures.
37.26 An international perspective.
37.27 Ability to act according to norms and rules of society.
37.28 Ability to participate actively in social activities.

A.4.2. Response Alternative

1. Not acquired at all
2. Poorly acquired
3. Acquired to some extent
4. Acquired well

Name:
Kazushi Okamoto

Affiliation:
Department of Informatics, Graduate School of Informatics and Engineering, The University of Electro-Communications

Address:
1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan

Brief Biographical History:
2002-2006 B.E., Kochi University of Technology, Japan
2006-2008 M.E., Kochi University of Technology, Japan
2008-2011 Dr. Eng., Tokyo Institute of Technology, Japan
2011-2015 Assistant Professor, Chiba University, Japan
2015- Assistant Professor, The University of Electro-Communications, Japan

Main Works:
- K. Okamoto, F. Dong, S. Yoshida, and K. Hirota, "Content-Based Image Retrieval via Combination of Similarity Measures," Journal of Advanced Computational Intelligence and Intelligent Informatics, 15(6), 687-697, 2011.
- K. Okamoto, K. Kawamoto, F. Dong, S. Yoshida, and K. Hirota, "An Evaluation Strategy for Visual Key Image Retrieval on Mobile Devices," Journal of Advanced Computational Intelligence and Intelligent Informatics, 16(5), 713-722, 2012.
- K. Okamoto, "Families of Triangular Norm-based Kernel Functions and their Application to Kernel k-Means," Journal of Advanced Computational Intelligence and Intelligent Informatics, 21(3), 534-542, 2017.

Membership in Academic Societies:
- Japan Society for Fuzzy Theory and Systems (SOFT)
- The Institute of Electronics, Information and Communication Engineers (IEICE)
- Information Processing Society of Japan (IPSJ)