New Orleans JAZZED about Engineering Education

Paper ID #17186

Bayesian Network Models for Student Knowledge Tracking in Large Classes

Mr. Chao Chen, Department of Computer Science and Engineering, University of South Carolina

Chao is a PhD student in the Department of Computer Science & Engineering at University of South Carolina. He is interested in applying machine learning algorithms and Bayesian statistics in social science study.

Mr. Ramin Madarshahian, University of South Carolina

Ramin is PhD student in Structural engineering in University of South Carolina. He also got Master of applied Science in Statistics at middle of his PhD program. His main focus of research is uncertainty quantification in engineering and scientific problems.

Dr. Juan M Caicedo, University of South Carolina

Dr. Caicedo is an associate professor at the Department of Civil and Environmental Engineering at the University of South Carolina. His research interests are in structural dynamics, model updating and engineering education. He received his B.S. in Civil Engineering from the Universidad del Valle in Colombia, South America, and his M.Sc. and D.Sc. from Washington University in St. Louis. Dr. Caicedo's teaching interests include the development of critical thinking in undergraduate and graduate education. More information about Dr. Caicedo's research can be found online at http://sdii.ce.sc.edu

Dr. Charles E. Pierce, University of South Carolina

Dr. Pierce is a Bell South Teaching Fellow and Associate Professor in the Department of Civil and Environmental Engineering at the University of South Carolina. He is a member of the American Concrete Institute, American Society of Civil Engineers, and American Society for Engineering Education.

Dr. Gabriel Terejanu, University of South Carolina

Gabriel Terejanu is an Assistant Professor in the Department of Computer Science and Engineering at University of South Carolina. Previously he was a Postdoctoral Fellow at the Institute for Computational Engineering and Sciences at University of Texas at Austin. He holds Ph.D. in Computer Science and Engineering from University at Buffalo. He is currently working on the development of a comprehensive uncertainty quantification framework to accelerate the scientific discovering process and decision-making under uncertainty. Some projects currently supported by NSF and VP for Research include discovery of novel catalytic materials for biorefinery industry, modeling and prediction of naturally occurring carcinogenic toxins, and development of statistical models for tracking individual student knowledge.

Bayesian Network Models for Student Knowledge Tracking in Large Classes

Chao Chen[1], Seyedramin Madarshahian[2], Juan Caicedo[2], Charles Pierce[2], Gabriel Terejanu[1*]
[1]Department of Computer Science and Engineering
[2]Department of Civil and Environmental Engineering
University of South Carolina, Columbia SC

Introduction

Arguably, the post secondary educational system is currently going through a major transition. On one end, the demand on Universities and colleges is growing while budgets are being reduced[1]. On the other hand, open access initiatives are making available a considerable amount of material to students and instructors[2,3]. This translates to higher demands on instructors with limited resources. This is of particular importance in a time when the cost of higher education has risen much faster than the average inflation[4]. In this landscape, the instructor is forced to optimize any available resources. One of the most important resources for instructors is time. Effective instructors not only help students in the learning process but also use meaningful evaluation strategies and provide targeted feedback to the student. However, providing good quality evaluation and feedback can become challenging, especially in large classes. In some cases, instructors might tend to over-test in an effort to give students feedback but the result could be overworked faculty and overloaded students. In other cases instructors might choose less assessment, depriving students of valuable feedback in the learning process.

The focus of this work is to use the scientific method to accelerate the assessment of student knowledge. The idea is that individual student knowledge at a certain point in time is nothing but a hypothesis/model that needs to be tested. Student modeling has been identified as ``the key to individualized knowledge-based instruction" [5,6]. Similarly, assessment instruments are nothing more than experiments to discover how well a student has mastered specific concepts. In this framework, experiments inform models and models guide experiments. It is this interplay between assessment instruments (experiments) and student knowledge models (hypotheses) that can accelerate the assessment of student learning to allow the instructor to timely intervene. This fits naturally with probabilistic methodologies such as Bayesian inference that formalizes the scientific method. Such a computational tool will go beyond grading, and it will allow instructors to provide formative feedback with respect to the challenging concepts specific to each individual student, suggest remedial interventions, and guide future examinations.

Statics is a common course for many engineering disciplines. This class is arguably the first class where students have to learn and apply an engineering way of thinking. Most students in science and math classes, before taking statics might have focused on identifying and using the appropriate equation to use for a particular problem. Statics is different. In this class students need to understand some basic concepts and learn how to express these concepts in equations to be able to solve problems. This shift in the way of thinking that students need to undertake is challenging for some students. The Statics Course (ECIV 200) at the Department of Civil and Environmental Engineering at the University of South Carolina is taught every semester. The

* Corresponding author: terejanu@cec.sc.edu

course is usually taken in the second semester of the freshman year and the estimated passing rate of the course is approximately 70%.

This work is based on the fundamental building block of Bayesian networks used to model and track student knowledge. It addresses an open fundamental problem in constructing and using knowledge models to assess learning, namely how to relate curricular structure to knowledge models and how to inform the models using assessment data. The data collected during ECIV 200 Statics - Fall 2015, has been used to identify the challenges in constructing and implementing these knowledge models. The main challenge identified during this study using data collected from 37 students over three separate quizzes, relates to instantiation of conditional probabilities of various answers given the knowledge of the concepts. This is because the conditional probability distributions for these cases when the concepts are not known are highly dependent on the question and the type of misconceptions that the students have at the time of assessment. To address this challenge we propose to learn these conditional probabilities directly from the data. Initial results based on parameter learning for our models are presented here for this pilot study conducted during Fall 2015.

Construction of Knowledge Student Models based on Bayesian Networks

Bayesian networks (BN) provide a simple graphical approach to intuitively implement student models and perform inference in the presence of uncertainty. BN is a theoretically sound framework for working with a probabilistic model that encodes the joint probability distribution over a set of random variables of interest. These variables can be divided in hidden variables and observed variables. Hidden variables cannot be directly observed, they can only be inferred from observations. In knowledge tracking, hidden variables consist of features used to assess the knowledge of the student and guide the remedial interventions. Hidden variables are concepts that may be known or unknown by a particular student (e.g. vectors in Fig.1). Observed variables are the answers to questions designed to test various concepts (e.g. *Quiz1_V* in Fig2).

In a BN, connections between variables are used to capture the causal process by which the observations are generated[7,8]. This representation however might be argued not to be suitable for knowledge tracking just by interpreting the causal links through positive influence, namely knowing concept vectors influences knowing concept vectors 2D. A better way to account for causality in this context is to embrace negative influence as well. Causal links should be used to describe also that not knowing concept vectors is a cause for not knowing concept vectors 2D. This approach is adopted here to construct conditional probabilities, which during inference facilitates the identification of challenging concepts.

With advances in computational power, inference in BNs of considerable size has become feasible, which at its turn has enabled the development of robust Bayesian tools such as GeNIe and SMILE (https://dslpitt.org/genie). This has opened the opportunity for the development of computerized tutoring systems, such as Bayesian Intelligent Tutoring System introduced by Burtz et al. (2006)[9,10], which has inspired the assessment tool proposed here. However, the slow adoption of intelligent tutoring systems in practice is attributed to the lack of best practices in constructing student models. This is due to a number of challenges that need to be addressed. How to select the concepts and questions to be included in the model? How should these features be connected? How should the models be initialized and updated?

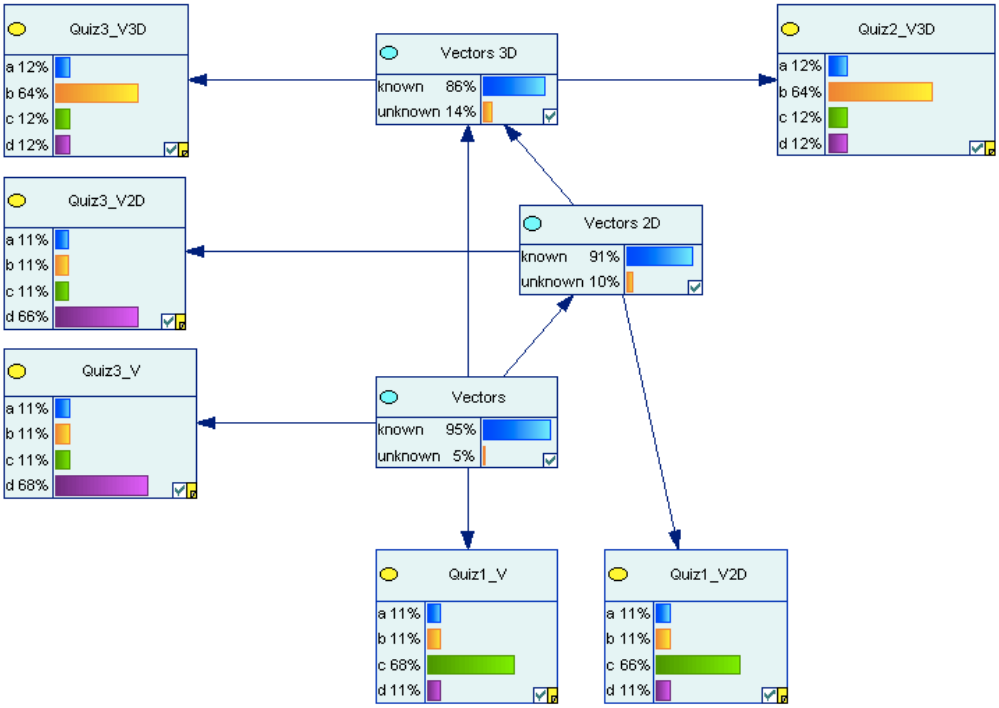

Figure 1 - Bayesian Network - Model Student Knowledge in Statics

Which concepts to include in the model and how to connect them? The concepts to be included in the knowledge model can be directly extracted from the curricular structure. Nonetheless, one needs to decide on the subset of concepts to be tracked. This decision should be taken based on how the information regarding the knowledge of these concepts will be incorporated in the instructional procedure. Here, besides the fundamental concept of vectors, we have also included concepts related to operations of vectors in 2D and 3D.

How to initialize these models? The initialization of the BN model is done by defining the prior probability distributions over all the nodes, which apparently seems to be a difficult task. Luckily, BN provide an intuitive way to factorize this joint distribution in marginal distributions over the nodes with no pre-requisites and conditional distributions between pairs of nodes. This factorization is intuitive because it accommodates ones understanding on how students progress in the class. Furthermore, the model is capable of representing prior knowledge on concepts studied in previous classes, and updating this knowledge based on new evidence.

Specifically, students in the Statics class should already possess knowledge of vectors from Linear Algebra class that is a pre-requisite to Statics. This expectation can be translated in a marginal prior probability for vectors, $P(Vectors=known) = 0.95$ and $P(Vectors=unknown) = 0.05$. With this knowledge the instructor can teach advanced concepts such as vectors 2D & 3D. The transition probability between pairs of concepts is defined by the capacity of the instructor to transfer that knowledge to the student. In our case, $P(Vectors2D=known|Vectors=known)=0.95$ and $P(Vectors2D=unknown|Vectors=known)=0.05$. In other words if the student knows vectors and the instructor is effective in teaching vectors 2D, then with high probability the student should know vectors 2D. Similarly, the statement $P(Vectors2D=known|Vectors=unknown)=0.05$

and *P(Vectors2D=unknown|Vectors=unknown)=0.95*, reads as follows: when the student does not know vectors, then with high probability this will impede him/her to understand vectors 2D. Similar conditional probabilities can be defined between all the linked concepts in the network. Note that as more concepts are introduced the probability that a student knows them is decreasing in the absence of testing. For example, the prior probability that vectors 3D is known is only 86% in the absence of any testing. However, the overall goal of this model is to infer from data whether these concepts are indeed known or not-known by each individual student. This is accomplished by using Bayes rule to calculate the posterior probability of concepts given the answers provided by the student to various questions.

Transition probabilities between concepts and questions are based on the type of answers provided. For example, a correct answer to *Quiz1_V* requires the student to master vectors (V), see Fig.2. This prompts *P(Quiz1_V=C|V=known)=0.7* and *P(Quiz1_V=X|V=known)=0.01* where X is either A, B, or D. This conditional probability corresponding to answering correctly can be easily constructed, however, the more interesting probability of answering incorrectly when the concept is not known is more challenging to define. This is related to the misconceptions that the students have at the time of assessment. Namely, we cannot distribute the probability equally among the answers if the vectors concept is not known, because some of the distractors are more likely than others. However, this is challenging to be determined a priori as compared with the conditional probabilities corresponding to knowing the concept. To address this challenge, we will take advantage of the parameter learning algorithms such as Expectation Maximization to adjust the conditional probabilities based on all the answers collected after each quiz. The advantage of this strategy is that it automates the construction of Bayesian networks and it provides valuable characterization for the proposed distractors, which can guide the instructor in designing remedial strategies.

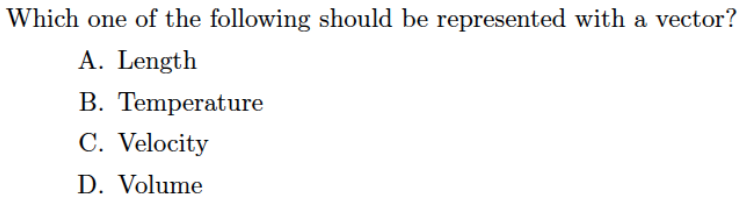

Figure 2 - Quiz1_V

One of the current computational challenging is that GeNIe/SMILE provides learning of the entire conditional probability distribution including the probabilities corresponding to answers given that the concept is known. While adjusting the probabilities conditioned of concepts not known is desired as argued above, optimizing the conditional probabilities given that the concept is known presents the risk of ending up with inconsistent logical relationships in the network for extreme datasets (e.g. when all students answer incorrectly). Currently, we are manually adjusting the conditional probabilities given that all the concepts are known to ensure logical relationships. Future work is planned to further constrain the parameter optimization.

Numerical Results from Pilot Study

The data collected from 37 students over three sequential quizzes has been used to inform the development of Bayesian networks for knowledge tracking. Each student will have his/her own

individual model. At the beginning of the class, prior to any testing the probabilities will be similar across models. However, once the testing begins these probabilities will be quite dissimilar from student to student, which in essence will provide a knowledge profile for each individual student. Three quizzes have been given during this study. The first quiz contains two questions related to vectors and vectors 2D, see Fig. 2 and Fig. 3. The second quiz contains just one question related to vectors 3D. And finally, the last quiz contains three questions corresponding to all three concepts. See Table 2 for the answers provided for the three quizzes.

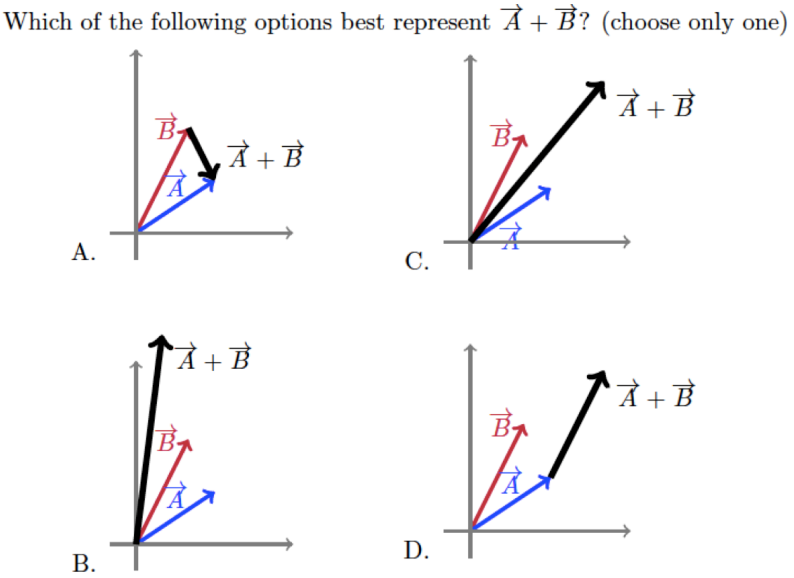

Figure 3 - Quiz1_V2D

The posterior probabilities can be an indication of the competency of the student in mastering a specific concept, see Table 3. These posterior probabilities can be compared against a specific threshold (i.e. 70%). Any drop below this threshold may trigger an intervention by the instructor and an action by the student. Note that tracking and making the probability of mastering a concept known to the student goes beyond the information provided by grading. Not answering correctly to a question right does not necessary prompt the student to revise the concepts associated with that question. We believe that specifically informing the student about his/her level of competency in various concepts and how this may impact his/her plan of study provides a much more direct and actionable information to the student. This particular hypothesis testing is planned to be tested in the following semesters. Before conducting any individual inference we need to adjust the conditional probabilities based on all the answers collected after each quiz using Expectation Maximization. The following results in Table 1 have been obtained after the optimization to learn the conditional probabilities using all the data from Quiz 1, which contains the two questions listed in Fig. 2 and Fig. 3.

Table 1: Conditional probabilities after maximum likelihood estimation

x	***P(Quiz1_V=x|Vectors=unknown)***	***P(Quiz1_V2d=x|Vectors=unkown)***
A	0.77	0.33
B	0.07	0.47
C	0.09	0.12
D	0.07	0.08

Table 2: Student answers for all three quizzes. The correct answer is provided in the header after the concept identifier V – Vectors, V2D – Vectors 2D, V3D – Vectors 3D

	Quiz 1		Quiz 2	Quiz 3		
ID	V/C	V2D/C	V3D/B	V/D	V2D/D	V3D/B
1			B	D	D	B
2	D	C	B	C	A	B
3			B			
4	C	C	B	D	C	B
5	C	C		D		B
6	C	D	B	D	D	D
7	A	A	A	D	D	B
8	C	C	B	D	D	B
9	C	C	B			
10	C	C	B	D	A	B
11	A	B	B	A	D	B
12	C	A				
13	C	C	B			
14	C	C	B	D	D	D
15	C	C	B	D	D	B
16	C	C	B			
17			A	D	D	B
18			B			
19	C	A				
20	C	C	B	D	D	B
21	A	A	B	D	A	A
22	C	C	B			
23	C	C	B	D	D	B
24	A	B	C	A	D	D
25	C	A	B	A	D	B
26	A	C	B	D	D	C
27	C	C	B	D	D	B
28	C	C	B	D	B	B
29	C	C	A	D	D	A
30	C	C	B	D	D	B
31	A	C	B			
32	C	C	B	D	D	A
33	C	C	B	D	D	B
34	C	C	B	A	D	A
35						
36	C	C	A	D	D	B
37	C	A	B	D	D	B

Table 3: The prior/posterior probability of knowing the three concepts V – Vectors, V2D – Vectors 2D, V3D – Vectors 3D after each quiz Q1, Q2, Q3.

	V=known [%]				V2D=known [%]				V3D=known [%]			
ID	Prior	Q1	Q2	Q3	Prior	Q1	Q2	Q3	Prior	Q1	Q2	Q3
1	95	95	96	100	91	91	93	99	86	86	90	96
2	95	98	96	9	91	98	96	9	86	92	92	15
3	95	95	96	96	91	91	93	93	86	86	90	90
4	95	100	100	100	91	99	99	99	86	94	96	96
5	95	100	100	100	91	99	99	99	86	94	94	95
6	95	99	98	100	91	88	80	98	86	84	78	96
7	95	60	58	96	91	55	52	96	86	54	51	91
8	95	100	100	100	91	99	99	100	86	94	96	97
9	95	100	100	100	91	99	99	99	86	94	96	96
10	95	100	100	100	91	99	99	91	86	94	96	89
11	95	15	16	40	91	3	4	29	86	7	10	34
12	95	99	99	99	91	89	89	89	86	85	85	85
13	95	100	100	100	91	99	99	99	86	94	96	96
14	95	100	100	100	91	99	99	100	86	94	96	98
15	95	100	100	100	91	99	99	100	86	94	96	97
16	95	100	100	100	91	99	99	99	86	94	96	96
17	95	95	95	100	91	91	90	99	86	86	85	95
18	95	95	99	99	91	91	95	95	86	86	92	92
19	95	99	99	99	91	89	89	89	86	85	85	85
20	95	100	100	100	91	99	99	100	86	94	96	97
21	95	60	66	36	91	55	62	9	86	54	61	11
22	95	100	100	100	91	99	99	99	86	94	96	96
23	95	100	100	100	91	99	99	100	86	94	96	97
24	95	15	30	41	91	3	1	8	86	7	1	6
25	95	99	99	100	91	89	92	99	86	85	89	96
26	95	93	88	97	91	94	90	99	86	88	85	91
27	95	100	100	100	91	99	99	100	86	94	96	97
28	95	100	100	100	91	99	99	99	86	94	96	96
29	95	100	100	100	91	99	99	100	86	94	94	88
30	95	100	100	100	91	99	99	100	86	94	96	97
31	95	93	95	95	91	94	96	96	86	88	91	91
32	95	100	100	100	91	99	99	100	86	94	96	92
33	95	100	100	100	91	99	99	100	86	94	96	97
34	95	100	100	100	91	99	99	100	86	94	96	92
35	95	95	95	95	91	91	91	91	86	86	86	86
36	95	100	100	100	91	99	99	100	86	94	94	95
37	95	99	99	100	91	89	92	99	86	85	89	96

Individual posterior probabilities are inferred after learning the parameters of the model after each quiz. These numerical experimentations have been used to uncover the challenges and further understand how these models perform in tracking student knowledge. Several cases can be pointed out for sanity check. For a student that has answered all the questions correctly (i.e. student #8 in Table 2 and 3), we can see that the posterior probabilities of knowing various concepts increase after each quiz. At the other extreme, we observe that the posterior probabilities drop significantly for someone that responded mostly incorrectly (i.e. student # 24).

There are several other cases of interest that have resulted due to the parameter learning. Namely, student #2 has responded incorrectly to the vectors question and correctly to vectors 2D questions in Quiz 1. Despite this, the posterior probability of knowing vectors has increased from 95% to 98%. The reason for this is that this particular student is the only one that has responded with D to the vectors question. As a result, D has not been identified as a powerful distractor as compared with A, see Table 1 and Fig. 2. Also note the difference after quiz 1 for student #2 and #31. In the absence of evidence from vectors 2D question, the posterior probability does indeed decrease however the correct answer for the vectors 2D question provides positive evidence for knowing vectors as well. Similarly, students #7 and #11, both have answered incorrectly to the two questions in quiz 1, however the decrease in their posterior probabilities for vectors and vectors 2D is dramatically different. The likelihood of these distractors has to be jointly explained. Namely, while there are exactly two students that responded (*Quiz1_V=A* and *Quiz1_V2D=A*) and (*Quiz1_V=A* and *Quiz1_V2D=B*), there are four others that have responded (*Quiz1_V=C* and *Quiz1_V2D=A*) and no others that have responded with B in the second question. As a result the likelihood of the distractor B in the second question is higher than A when the vectors 2D is not known.

While there are 12 students that have responded incorrectly to at least one question in quiz 1, not all of them need remedial intervention. Based on posterior probabilities and a competency level of 70% only four students are identified as needing an intervention at this point (students #7, #11, #21, #24). The performance of the last three students remains below the competency level after the third quiz, which provides some positive evidence for the proposed approach in tracking student knowledge. On the other hand, the performance of student #7 has significantly increased above the competency level after the last quiz. Student #2 has not been identified after quiz 1 as someone needing a remedial intervention, however the evidence from the last quiz points to a significant decrease in concept knowledge. These results are not currently at the level to claim any statistical significance, however this pilot study has provided valuable information regarding the construction of Bayesian models and it has set the plan to independently evaluate the results of the model in the following semesters. The proposed evaluation is to significantly expand the model with more concepts and questions and track the knowledge of the students over a larger period of time. A random number of students will be selected from the best and worst performance categories as indicated by the model, and they will be independently interviewed by faculty members to confirm whether they know or not know various concepts.

Conclusions

We have presented a methodology for constructing and using knowledge models to assess student knowledge in large engineering classes. This methodology has been informed by an initial pilot study conducted in the Fall 2015 Statics comprising data collected from 37 students

over three sequential quizzes. The main challenge identified during our first pilot study was related to the instantiation of the conditional probabilities of the distractors given that the concepts are unknown. The reason for this is that the probability distribution for these cases is highly dependent on the question and the type of misconceptions that the students have at the time of assessment. To address this challenge, we have used Expectation Maximization to adjust the conditional probabilities based on the all answers collected after each quiz. The advantage of this strategy is that it automates the construction of Bayesian networks and it provides valuable characterization for the proposed distractors, which can guide the instructor in designing remedial strategies. The numerical results provide some anecdotal positive evidence for the feasibility of the proposed approach in tracking student knowledge. Future evaluations based on independent interviews are planned to assess the performance of the model.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. 1504728.

References

1 Christensen, Clayton M and Eyring, Henry J (2011) *The innovative university: Changing the DNA of higher education from the inside out*, John Wiley & Sons.

2 Daniel, John (2012) 'Making sense of MOOCs: Musings in a maze of myth, paradox and possibility'. *Journal of Interactive Media in Education*, 3.

3 Brown, John Seely and Adler, Richard P (2008) 'Open education, the long tail, and learning 2.0'. *Educause review*, 43(1), pp. 16–20.

4 Archibald, Robert B and Feldman, David H (2010) *Why does college cost so much?* Oxford University Press.

5 Mayo, Michael and Mitrovic, Antonija (2001) 'Optimising ITS behaviour with Bayesian networks and decision theory'. *International Journal of Artificial Intelligence and Education*, 12, pp. 124–153.

6 Greer, Jim E and McCalla, Gordon I (2013) *Student Modelling: The Key to Individualized Knowledge-Based Instruction*, Springer Science & Business Media.

7 Yoder, Brian L (2012) 'Engineering by the Numbers'. *American Society for Engineering Education, Washington, DC. http://www.asee.org/papers-and-publications/publications/collegeprofiles/2011-profile-engineering-statistics.pdf*.

8 Pearl, Judea (2014) *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann.

9 Prince, Michael, Borrego, Maura, Henderson, Charles, Cutler, Stephanie and Froyd, Jeff (2013) 'Use of research-based instructional strategies in core chemical engineering courses'. *Chemical Engineering Education*, 47(1), pp. 27–37.

10 Butz, Cory J, Hua, Shan and Maguire, R Brien (2006) 'A web-based Bayesian intelligent tutoring system for computer programming'. *Web Intelligence and Agent Systems*, 4(1), pp. 77–97.