Intelligent Tutoring System using Computerised Adaptive Testing and interaction logs for MOOCs

M.Tech Project Report

Submitted by

Nilesh Birari
Roll No : 123059015

under the guidance of

Prof. D. B. Phatak

Department of Computer Science and Engineering
Indian Institute of Technology, Bombay

September, 2014

Contents

List of Figures

Abstract

A massive open online course(MOOCs) is an online course aimed at unlimited participation and open access via the web. But these courses does not provides personalised attention for each student like in classroom learning. ITS fill this gap by providing personalised guidance for every student with different learning need.

In this paper we will see the different adaptive and intelligent technologies and how these techniques can be used for building ITS. ITS aims to provide personalised instruction for each student, so the student model is the key component of ITS which store the student information(student knowledge state) . Diagnosis of the student knowledge state is the most difficult process due to uncertain and imprecise information about the student. To manage uncertainty in data Bayesian network has the most sound mathematical foundations and also a natural way of representing uncertainty using probabilities. The proposed system, which uses adaptive technologies for personalised guidance for each individual using state of knowledge represented in Bayesian student model.

Chapter 1

Introduction

1.1 Challenges in MOOCs

A massive open online course(MOOCs) is an online course aimed at unlimited participation and open access via the web [1]. MOOCs offers wide variety of courses to the massive number of students. Benefits of such on-line courses is they are classroom independent and platform independent.The courses started at one places can be used by thousands of learners all over the world. But the courses offered are nothing but the static traditional hypermedia of the various learning material. Each student registered in MOOCs has different characteristics such as knowledge level, preferences, interest , goals, cognitive style ,learning style, prerequesite knowledge etc. The courses offered by MOOCs could not able to meet the need of each and every individual with different learning demands. Due to which most of the student not able to achieve satisfactory learning in course.

Figure 1.1 shows the learning workflow in MOOCs. When a new course start, instructor

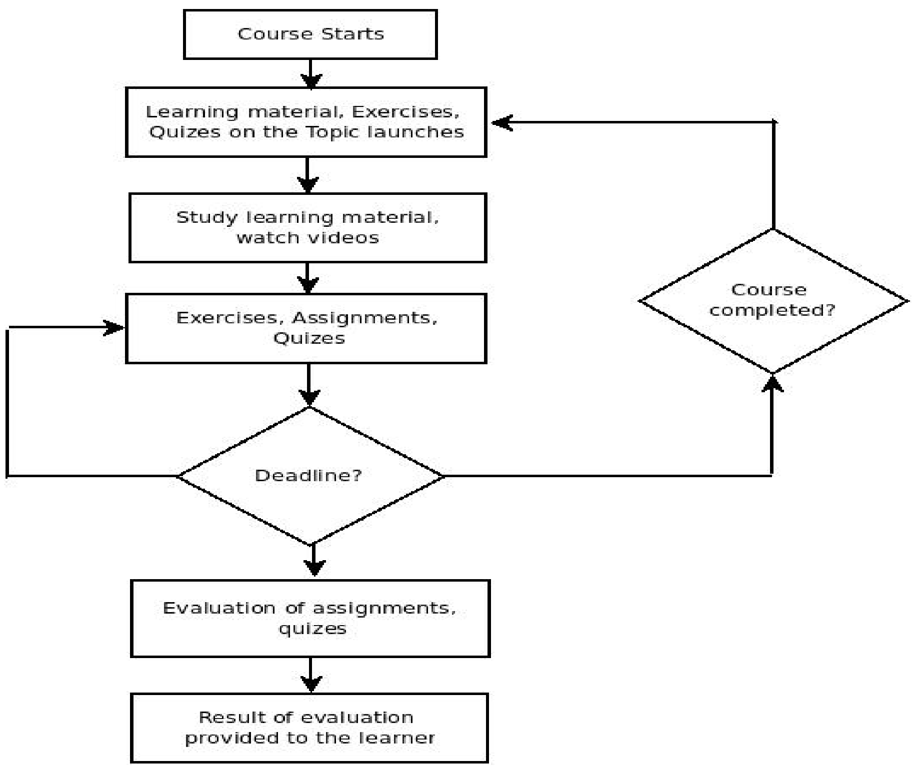

Figure 1.1: Work flow of MOOCs

puts learning material on topics. As time passes during the subsequent assessment learner goes through the available learning material. learning material consist of Videos, audio, text etc. Going through learning material learner attempt for exercises, assignments, quizzes, exams etc. After assessment results evaluated are provided to the learner and same procedure follows until the course finishes.

For the diverse learner population, this system will suffer from inability to be "all things to all people" . The system will present the same static explanation and suggest the same next page to learner with widely different learning characteristics [2].

Due to the MOOCs assessment and evaluation process for the course, it is not possible to detect weather student understand all the concept covered in course. The evaluation system allows to pass the student even if he can perform good in major of the concept. So there is strong need of the system which can identify the concept in which student not perform well along-with providing some personalised learning material for those concepts.

In traditional classroom based courses, Due to limited number of students, teacher can give special attention for each and every student for their classroom activities and can observe learning style, knowledge level, interests, preferences etc. In such environment teacher can diagnose the student knowledge on various concept by asking questions, and teach on the basis of the diagnosis result for each student. But Due to development in technologies and need to serve for the large number of population with courses are offered on web with open access to all. The number of student participating increases rapidly and a single teacher can not take care about each and every student learning need. So to take care of such an individuals a technologies like Intelligent Tutoring System are emerged.

1.2 Intelligent Tutoring System for MOOCs

Intelligent and Adaptive technologies proposes a solution to overcome limitations of the traditional online courses. ITS systems build a model of the knowledge of each individual learner, and use this model throughout the interaction, in order to adapt to the needs of individual learner. System offers personalised learning material as per individual learning need.

In this work we will propose a solution for the problems mentioned above. The proposed solution built a model for each individual, during the assessment and interaction with the system. The model represent learners knowledge level. Using the student responses for the assessment, system estimate performance of each student as for topic and able to make prediction whether student understand each concept/topic he learned. Using this model, system analyse individual student need can provide personalised learning material as per individuals demand.

1.3 Intelligent Tutoring System Components

Majority of the ITS systems build has modularised in some components according to functionality. To build ITS for the MOOCs, we need to study some of the important component. As illustrated in figure 1.2 ITS contains following components.[3] [4]

- Domain Model
- Student Model
- Tutoring Model
- Interface Model

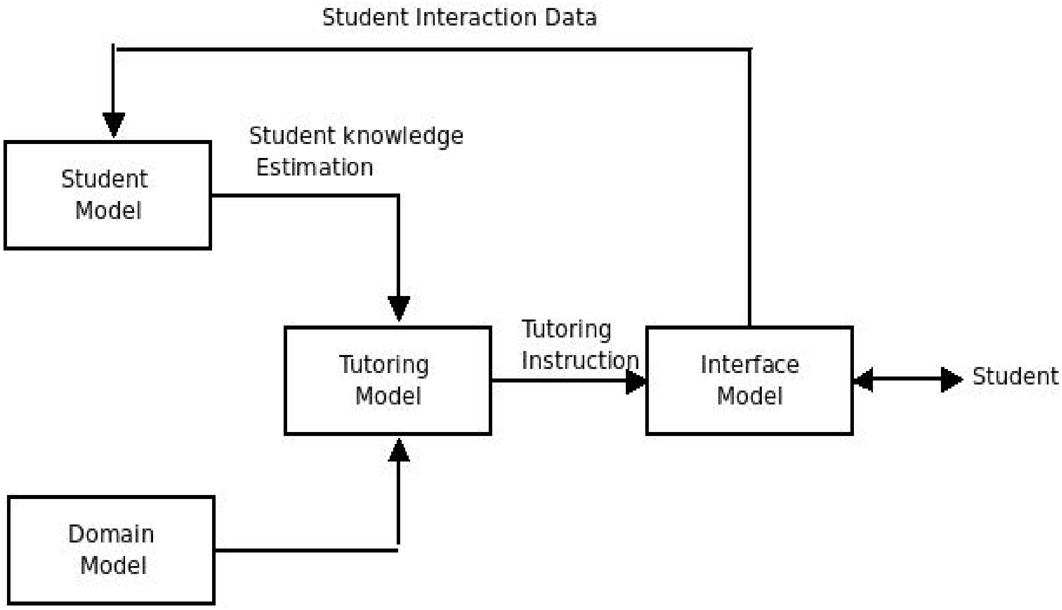

Figure 1.2: Components of ITS

Domain Model- Domain model also known as expert model, it stores information about the material being taught. This model contains concepts and relationship between concepts, solutions for the question, lessons and tutorials, and problem solving strategies of the domain. It stands as a backbone of the content of the system.

Student Model- It represents domain dependent and independent characteristics of each user. These characteristics are acquired from user explicit responses, analysis from the interaction data, through assessments etc. It consider as a core component of the ITS system.

Tutoring Model- To make the choices about tutoring and action, it accepts information from domain model and student model. it guides learner with respective their current state of knowledge to reach proficiency with the target skill.

Interaction Model- The model concerned with the representation of the user(user model) and representation of the application(domain model). the main aim of the model is capturing the appropriate raw data of the learner and representing the interface.

1.4 Summary

Due advancement in technologies and increasing demand of the users, wide range of courses are offered on web. MOOCs offers such online courses with free access to all, so

the large number of users are participating in such courses, each with different learning need, but unlike conventional classroom coaching it is not possible to pay special attention for each and every single student in such massive open courses. So this brings up the need of the Intelligent tutoring system. ITS builds a model for each individual and use this model throughout the interaction, in order to adapt to the needs of individual learner.
‘

In next section we will study Adaptive Hypermedia, In Order to adapt changes for each individual. In which we will see the the methods and technologies employed to make the system adaptive. Followed by will study the methods and technologies used to implement Intelligent Tutoring System and most important the student model which used to store the student information of current state of the knowledge for each individual. Due to uncertain and imprecise information student modeling is the difficult task, so we will see how this is handled by the use of Bayesian network for student modeling. Finlay we will propose a solution with the study of these things.

Chapter 2

Adaptive and Intelligent Technologies

As we have seen the benefits of MOOCs and need of the ITS for MOOCs. In this section we will take a brief survey of the existing ITS and Adaptive Hypermedia Systems. Along with we will study what are the adaptive and intelligent technologies available, and how they are implemented on web. Most of the adaptive and intelligent technologies are inherited from Adaptive Hypermedia and Intelligent Tutoring System.

2.1 Adaptive Hypermedia

Static hypermedia applications has some limitation that they provides same set of presentation and links to all user. For example they provides same static explanation and same next page to students with widely different educational goals and knowledge of the subject.
To overcome the limitations of static hypermedia, adaptive hypermedia system builds model of knowledge, goals and preferences of each individual student and use this throughout the the interaction with student, in order to adapt to the need of that student. For example, a student in an adaptive educational hypermedia system will be given a presentation that is adapted to his knowledge of the subject, and a suggested set of most relevant links to proceed further.[5]

What features of user can be taken into account when providing adaptation? There are various features of the user different for different users, out of adaptive system considers those are related to the current context of the user work and useful for the adaptation. Adaptive system use the features include user knowledge, goals, preferences, background, hyperspace experience, user interests and individual traits etc. [2]

AHAM is one of the most popular reference model to support adaptive hypermedia authoring. AHA(Adaptive Hypermedia Architecture) system which is authoring tool for the adaptive hypermedia application which is build on the same reference model. To build Adaptive Hypermedia Application one should familiar with such models. So in next section we will see the framework to build adaptive hypermedia and followed by will see the method's and techniques of adaptive hypermedia.

2.1.1 A Framework for Adaptive Hypermedia

The knowledge contain in the hypermedia document is present in the form of concepts and concepts are connected through concept relationship. [6]
Adaptive Hypermedia performs three functions

- When user interact with the system, system captures interaction with the user. Based on this observation Hypermedia system maintains model of the user about each domain concept. System store a knowledge value for each concept. Knowledge value shows the information about how much user does knows about the concept and also represent whether user read that concept or not.

- According to current knowledge, interest or goal nodes or pages are classified into into several groups using user model. The system manipulated link anchors within pages (and link destination) to guide user towards interesting, relevant information. Link anchor could be specially annotated, disabled, or removed. This method called as a adaptive navigation support.

- Content of the pages contains appropriate information i.e. expected level of difficulty or details for each user according to user model. When presenting system conditionally show, hide, highlight, or dim conditional fragments on a page. This is done in order to ensure that page contains appropriate information for a user at given time. This method called as a adaptive presentation.

2.1.2 Adaptive Presentation

Adaptive presentation [6] of the information is nothing but the manipulation of the text fragments. this manipulation can be:

- Providing prerequisite, additional, or comparative explanation: Additional information provided for the user with specific knowledge state. Additional information contains prerequisite knowledge, additional details, or a comparison with a previously known concept. Techniques used for such presentation are Conditional inclusion of fragments and Strechtext. In Conditional inclusion of fragments, user model and concept relation model(domain model) determines which fragment should be displayed. Conditional inclusion of fragments is used in the AHA system. In Stretchtext, for each fragment there is a placeholder. System determines which fragments to be "stretched" or "shrunk" (i.e. only the placeholder is shown). This only determines initial presentation, user can stretch or shrink fragments through mouse click. System captures user action after few mouse click to better predict for next subsequent pages request.

- Providing explanation variants: The same information can be presented in different ways, depends on level of difficulty value in user model. Explanation variants are used in Anatom-Tutor, Hypadapter.

- Reordering information: Some user may prefer example before definition, while other prefer other way around, all this reordering is done using the user model values.

2.1.3 Adaptive Navigation Support

The manipulation [6] of links that are presented within nodes (pages) is typically done in one or more of the following ways:

- Direct guidance: System informs student that 'next' button lead to the most appropriate page in hyperspace according to the current knowledge level.
- Sorting of links: The presented list of links is sorted from most relevant to least relevant.
- Link annotation: Link anchors are presented differently depending on the relevance of the destination.
- Link hiding: Inappropriate or non-relevant information containing link are hidden from presented page.
- Link disabling: Inappropriate links are disabled.
- Link removal: Inappropriate links are simply removed.

2.2 ITS technologies

ITS [7] uses knowledge about domain, student and teaching strategies to support individualised learning and tutoring. Curriculum sequencing, problem solving support are the core technologies used by the most of the ITS.

2.2.1 Curriculum sequencing

Curriculum sequencing[7] finds 'optimal path' through the learning material by providing student with the most suitable individually sequence of knowledge units to learn and sequence of learning tasks (examples, questions, problems, etc.). There are two kinds of sequencing include active and passive. Active sequencing build the individual path to achieve the goal. Passive sequencing does not require any learning goal, it start when user unable to solve some problem or user requires some help. The goal of curriculum sequencing is to offer student appropriate learning material to fill the gap between student knowledge to resolve misconception. Most of the existing systems guides students to the fixed learning goal i.e. whole subject. There are two level of sequencing possible are High-level sequencing or knowledge sequencing and Low-level sequencing or task sequencing. High-level sequencing determines next learning subgoal i.e. next concept, set of concepts, topic, or lesson to be taught. Low-level sequencing determines next learning task within current subgoal, learning task contains problem, example, test etc. in the same subgoal.

Sequencing system is mostly driven by the difference in student knowledge and domain knowledge. System can also use of students preferences on the type of media of available learning material to drive sequence of task within a topic. Few system exist they posses only kind of learning material is questions. SIETTE is an example of a Web-based adaptive testing system. They generate adaptive sequence of questions to assess students knowledge. So this kind of system are incomplete and they have to be used as a component in ITS.

2.2.2 Problem solving support

problem solving support [7]is one of the most widely used technologies in most of the ITS systems. There are three different approaches are Intelligent analysis of student solutions, Interactive problem solving support, Example-based problem solving support. All these technologies uses different way to help student in problem solving process.

- *Intelligent analysis of student solutions* technology decides whether solution is correct or not, find out what is wrong or incorrect in the solution, and identifies which missing or incorrect knowledge may responsible for the incorrect solution. Intelligent analyzers provide the student error feedback of the incorrect problem solution and update the student model.

- *Interactive problem solving support* provides student with intelligent help on each step of problem solving. Help is in the form of signaling about wrong step, hints for current step etc. System monitors student action, understand them and use this understanding to update student model and provide them appropriate help.

- *Example-based problem solving* technology help students to solve problem by suggesting them relevant successful problem from their earlier experience.

2.3 Existing Systems

In the field of adaptive and intelligent technologies for education, there are many systems were proposed from last two decades. Most of the system build uses the techniques listed above. The table shows the some of the most popular systems and the technologies used by those systems. [8][9] [7]

System	Description	Technologies used
AHA	open adaptive hypermedia architecture.Platform for adaptive hypermedia	adaptive presentation, adaptive navigation support
ELM-ART	Intelligent interactive system to learning programing in lisp	adaptive sequencing, adaptive presentation, adaptive navigation support,problem solving support
InterBook	Authoring and Delivering adaptive electronic textbooks	adaptive sequencing, adaptive presentation, adaptive navigation support
KBS Hyperbook	textbook in hypertext document	adaptive sequencing, adaptive navigation support
NetCoach	Authoring system that meet the needs to create adaptive learning course	curriculum sequencing, adaptive annotation of links
Metalink	Authoring tools for adaptive hyperbook	page sequencing, adaptive presentation
SIETTE	ITS for classification and identification of different vegetable species	Question sequencing
SQL-Tutor	support learning SQL	Intelligent solution analysis

2.4 conclusion

Most of the existing adaptive and intelligent systems are aimed at specific applications i.e. specific to some domain. Such systems are closed and difficult to reuse for other domain application. Few systems like AHA offers authoring tools to build the adaptive course content for each individual. But with the study of MOOCs we can conclude that there is requirement of reliable platform which can make teaching and learning easier. The platform is nothing but the Learning Management System(LMS) that can handle serving tests, grading assignments, pushing course material, providing user friendly interface, communication through chat rooms, discussion forum and many other feature present in current LMS. So these features difficult to build and handle by the any stand alone adaptive learning system. Proposed solution for this problem is to integrate the adaptive and intelligent technologies with the existing LMS.

Chapter 3

Student Modeling

The goal of an ITS is to adapt instruction as per individual knowledge level for each student. Student model is the key component that allows personalised instruction to be adapted to each student. Student model contains all information about student current state of knowledge that allows ITS to provide the personalised tutoring. [10] An ITS collect data from various sources for creating and maintaining student model. Sources include observing student interaction, quizzes and exams, or from explicitly request from user.

In next section we will see the features used for student modeling and some of the popular modeling techniques. Followed by will see the uncertainty management in student modeling.

3.1 Information in Student model

Student model contains important information about student such as domain knowledge, preferences, interest, goal, tasks, learning style, background etc. Content of the student model can be divided into domain specific and domain independent information. [11]

Domain Independent Information contains learner goals, interests, background and experience, individual traits, aptitudes and demographic information.

Domain specific information shows the status and degree of knowledge and skills student achieved in certain domain. It is organized as in the form of knowledge model that has many elements which student needs to learn. User knowledge is the most commonly used feature in most of the adaptive education systems for student modeling. some of the models used to represent knowledge of the domain are vector model, overlay model, and fault model.[12]

- *Vector model*- is the simplest form of knowledge representation. The vector consists of concepts, topics or subjects in domain. Each element of the of the vector is a real number or integer represents degree which learner gains knowledge about that concepts, topics or subjects.

- *Overlay model*- To represent an individual user's knowledge as a subset of the domain model, is the basic Idea of overlay knowledge modeling. In domain model each element represents a concept, subject or topic. So, the structure of user model is

constructed from the structure of domain model. Each element in user model has a specific value measuring the user's knowledge about that element corresponding to each element in domain model. This value is considered as the mastery of domain element and represented in the form of binary (knows or ignores), qualitative (good, average, weak, etc.) or quantitative (the probability of knowing or not, a real value between 0 and 1, etc.).
Overlay modeling approach is based on domain models, which is constructed as knowledge network. The complexity of the model depends on the Domain Model structure and on the estimate of the student knowledge, which is acquired through assessments and analysis of the student's interaction with the system. Fig 3.1 represent simple overlay model in which number indicates mastery of the concept on the scale of 10.

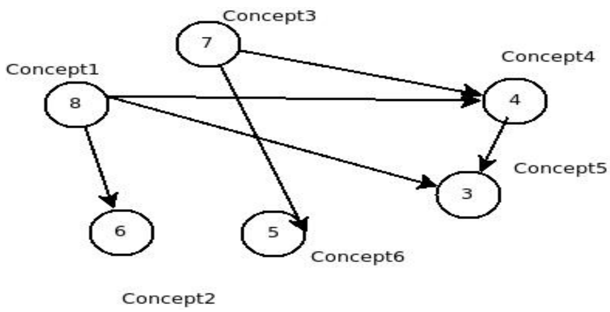

Figure 3.1: Overlay model

- *Fault model*- Unlike vector model or overlay model fault model is used to describe the lack of learner's knowledge. The model contains learners errors or bugs. Information from fault model, can be used to deliver learning material, concepts, subjects or topics that users don't know.

3.2 Uncertainty-Based User Modeling

The state of knowledge is generated from student behavior during interaction with the system, i.e. inferred from information available for each student. Diagnosis is the process that infers student knowledge state from the available student information. Diagnosis is the most complicated process in an ITS. Due to uncertain (we are not sure that available information is absolutely true) and/or imprecise information(the values are not completely defined), treatment of inferring knowledge state from such information is difficult process. Suppose for example student fail to answer question , so most probably student doesn't know the concept is uncertain information, and observation of student reading about concept is imprecise observation to estimate student knowledge.[10]

Artificial Intelligence has addressed this problem in various ways such as rule-based systems, fuzzy logic, Bayesian Networks etc. Bayesian networks are the most powerful approach for uncertainty management in Artificial Intelligence. This technique combines the rigorous probabilistic formalism with a graphical representation and efficient inference mechanisms. [13] Due to sound mathematical foundations and natural way of representing uncertainty using probabilities, Bayesian networks (BNs) used by many system designer for uncertainty management.

Chapter 4

Bayesian Diagnostic Algorithm for Student Modeling

In this section, we will see approaches to implement student model for more accuracy in diagnosis of student knowledge at every conceptual level. The student model defined is an overlay student model , that is, the student's knowledge is considered as a subset of the expert's knowledge. Described in previous chapter. We will describe the structural model used in our approach to Bayesian student modeling. For defining model we have proposed two different approaches addresses to different problems in diagnosis. First approach is based on Bayesian Network and application of Computerized Adaptive Tests to improve efficiency and accuracy of the diagnosis process. [14] Another approach look at the problem when adaptive testing environment is not available. This approach proposed the solution to perform diagnosis with available student data. In such situation diagnosis process will be more difficult due to very less number of available evidence for each concept. On the basis of these approaches we will see the combined approach for efficient student modeling and its implementation to improve the diagnosis process.

In next sections we will see theoretical background of Bayesian network and Computerised Adaptive testing, and approaches to handle diagnosis using Bayesian network.

4.1 Bayesian Network

A Bayesian Network [15]is directed acyclic graph in which nodes in graph represents variables and arcs between nodes represents probabilistic dependency among variables. The parameters used to represent uncertainty are the conditional probabilities of node given its parents, if the variables of the network are $\{X_i, i = 1, ...N\}$ and parents of the node X_i represented by $Pa(X_i)$ then the the parameters of the network are the conditional probability distributions $\{P\left(X_i/Pa\left(X_i\right)\right), i = 1, ..., n\}$ for each variable given it's parent(for the nodes without parents, the a prior distributions). This conditional probabilities describes joint probability distribution of the network, distribution for the entire network as,

$$P\left(X_1, ...X_n\right) = \prod_{i=1}^{n} P\left(X_i/Pa\left(X_i\right)\right)$$

To define Bayesian Network need to specify following things

- The set of variables $X_1, X_2, ..., X_n$
- The set of relationship(arc) among those variables. The arcs represent casual influence among the variables, and network form with these arc and variables should not contain any cycle.
- For each X_i the conditional distribution given its parent is $P\left(X_i/Pa\left(X_i\right)\right), i = 1, ..., n$

Once a BN is created, it can be used to make inferences about the values of the variables in the network. Bayesian propagation algorithms make such inferences using available information, usually set of observations or evidences. This process is sometimes referred to as beliefs update. After beliefs update, a posterior probability distribution reflects the influence of evidence for each associated variables. Inference are made for diagnostic or predictive. Diagnosis is for identifying most likely cause given a set of evidence. Evidence in that context can be referred a symptoms or fault. On the other hand, Prediction is made to identify the most likely event occurrence given a set of observations. In a BN, every variable can be either a source of information (if its value is observed) or object of inference (given the set of values that other variables in the network have taken). [15]

We are using Bayesian Network to define student model , in which variable are used to represent different things depending on the domain. The variables can be used to represent rules, concepts, problems, abilities, skills, etc. The link between variables represents relationship like part-of, prerequisite-of etc. After that conditional probabilities must be specified. Once everything is setup Bayesian Network can be used to make inference about the values of the variables in the network. Observations or evidences are used as a available information to make the inferences using probability theory.

4.2 Computerised Adaptive Tests

Computerised adaptive tests(CAT) [16] is test administered by a computer in which selection of next question ask and decision to stop the test are performed dynamically by the computer based on a student profile created and updated with response of the student for asked questions. The main difference between CATs and traditional tests is the capability to adapt to each individual student. In this way, a CAT allows to evaluate like human teacher by asking him/her a few selected questions that were progressively more difficult for students with higher levels, or easier for students with lower levels . The great advantage is significant decrease in test length with better estimation of student knowledge level. The adaptive question selection algorithm chooses best(most informative) question to ask next, with the current estimation of the student knowledge.

Significance of using CAT for ITS implementation is to evaluate knowledge of student in domain along with providing platform for problem solving support. Aim of the ITS to provide the problem solving support to the learner, with the use of CAT it can achieve by proposing challenging questions to the learner. Unlike traditional problem solving platforms in which problems were selected not by considering learner ability in mind. Due to learner with higher abilities bored by solving easy questions and learners with lower abilities get frustrated or loses their interest with difficult questions. CAT proposes challenging question by tailoring the ability of learner to the question difficulty level, that help to maintain learner interest in problem solving and through which he/she gains knowledge.

CAT is an iterative process that start with initial estimation of the student proficiency and has the following steps:

1. The questions from the database (that have not yet selected) are examined to determine which will be the best to ask next that will give maximum information according to the current knowledge estimation.

2. Examinee responds to the question asked.

3. from the response given new proficiency level is computed.

4. Until stopping criteria met steps 1 to 3 are repeated.

This procedure is illustrated in Figure 4.1 :

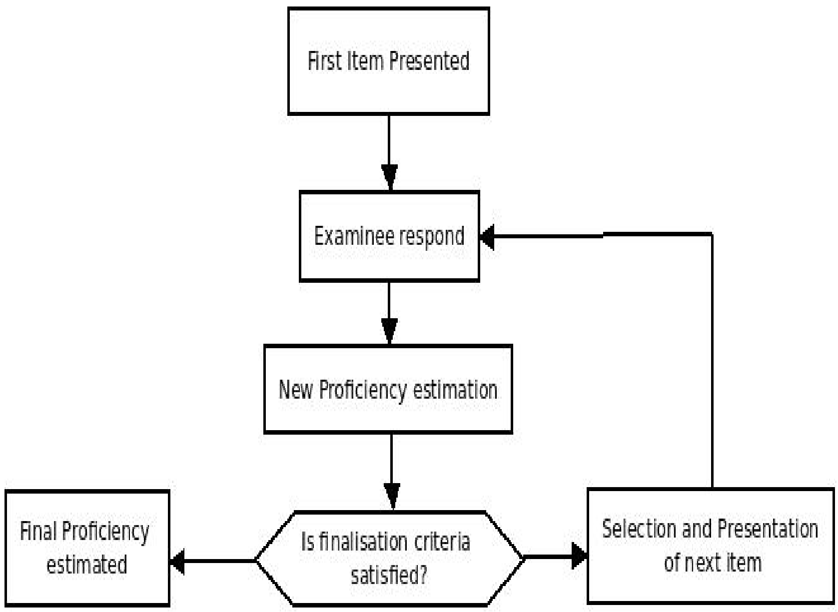

Figure 4.1: Flow diagram of an adaptive test

The elements in the development of a CAT are:

- *Item Response model.*This model is associated with each question. This model describes how examinees answer the item depending on their level of ability. The result obtain during measuring the proficiency be invariant of type of test and individuals that takes test.

- *Item pool.* A good item pool contain a large number of correctly calibrated items at each ability level. To achieve better result it must required to have better quality of item pool.

- *Ability estimation.* a method to compute the student's ability level according to response to the question asked.

- *Input proficiency level.* Suitably choosing the difficulty of the first item can considerably reduce the length of the test. Different criteria can be used like taking the average level of knowledge of the examinees that have taken the test previously.
- *Item selection method.* Adaptive tests select the next item to be asked depending on the proficiency level of the examinee that is obtained from the answers to items previously asked. Selecting the best item to ask which can improve accuracy and reduce test length.
- *termination criterion.* There can different termination criteria are used to decide when the the test should finish. These criteria includes target measurement precision has achieved, fixed number of items presented, time finished etc.

4.3 Integrated Approach to Bayesian Student Modeling using adaptive testing

Bayesian Network is build using the overlay modeling approach in which knowledge node in the network corresponds to the concept in domain model. In this section we will see the nodes defined for the Bayesian student modeling and relationship between nodes. Specifying conditional probabilities is the most difficult task in defining Bayesian network. The model simplify the issue of specifying cognitional probabilities.[17]

In order to use the Bayesian Network as a basis to perform adaptive testing, needs to define a model in which nodes and link between nodes should be identified. The nodes considered are:

4.3.1 Nodes and relationship among nodes

To define Bayesian student model the nodes considered are:

- *Evidential nodes*: evidential nodes are the test used in adaptive testing that can be answered correctly/incorrectly, which is represented by E. These nodes are used to collect the information relevant to the student's knowledge state.
- *Knowledge nodes*:Which are defined at three different levels of granularity which consist of concepts, topics, and subject are represented by C, T, and A respectively. Different level of granularity provides detailed information about the student knowledge level, as required by the ITS. This kind of structure enables system to understand exactly which part of the subject mastered/not mastered by the student.
- *Auxiliary nodes*: These nodes are used for modeling convenience only. For simplifying the process of specifying conditional probabilities between concept and evidence, and item selection in adaptive testing. each auxiliary node is associated with only one test item. Auxiliary nodes are used, which are represented by B in network.

Relationship among those nodes are:

- *Aggregation relationship*: As mentioned above each knowledge node has a conditional probability distribution given its parent. The aggregation is exist only between knowledge nodes in granularity hierarchy.

- *Relationship among concept and auxiliary nodes*: It is just like aggregation relationship, exist between concept and auxiliary nodes. It represent influence of the related concept weight on which the relation is defined.

- *relation between auxiliary nodes and test item* Mastering/ not mastering the node has causal influence in answering related questions correctly. In this relationship auxiliary node represent mastering of the associated concepts.

The Bayesian Network defined is depicted in Figure 4.2. It is divided into two parts, which overlaps in concept nodes:

- Diagnosis process is performed by a part which contains concept nodes and test items. At this stage the set of concepts mastered by the student are inferred from student's answers to the test items.

- Evaluation process contains knowledge nodes, In which from the probability of knowing each concept(obtained in previous stage), the state of knowledge of each topic estimated. And from state of knowledge of topics, the knowledge of subject is estimated.

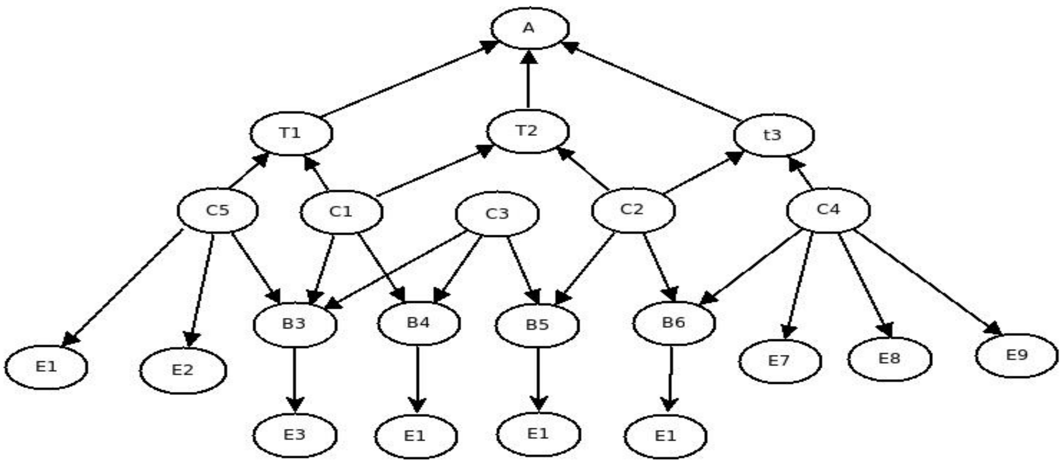

Figure 4.2: Bayesian Network for Adaptive Testing

4.3.2 Parameter specification

Once the model has been defined , need to specified required parameters. It is one of the most difficult problem for using Bayesian Network. Next we will see the proposed approaches for the parameter specification.

1. The prior probability of knowing the concept. Prior probability of mastering the concept can be specified from the information available about the particular student, if information is not available the uniform distribution is used. Information consist of accessing related material on concepts, time spend etc.

2. Conditional probabilities of each knowledge node given its parents. Conditional probabilities are computed on the basis of importance of each knowledge node in aggregated knowledge node. The importance of each knowledge node is decided by weight assigned to that node.

- Let $\{C_{ij}, i = 1, ...n_j\}$ be the set of related concept in topic T_j, and w_{ij} represent the importance of concept C_i in topic $T_j, i = 1, ...n_j$. Then, for each $S \subseteq \{i = 1, ...n_j\}$ required conditional probability distribution is given by:

$$P\left(T_j / \left(\{C_{ij} = 1\}_{i \in S}, \{C_{ij} = 0\}_{i \notin S}\right)\right) = \frac{\sum_{i \in S} w_{ij}}{\sum_{i=1}^{n_j} w_{ij}}$$

- Let $\{T_i, i = 1, ...s\}$ be the set of related topics in subject A, and α_i represent the importance of topic T_i in subject A. Then, for each $S \subseteq \{i = 1, ...s\}$ required conditional probability distribution is given by:

$$P\left(A / \left(\{T_i = 1\}_{i \in S}, \{T_i = 0\}_{i \notin S}\right)\right) = \frac{\sum_{i \in S} \alpha_i}{\sum_{i=1}^{S} \alpha_i}$$

Aggregation relationships are used to obtain information about how well student mastered topic or subject. By using this network it is possible to obtain estimation of subject proficiency or understanding of each topic, and this information allows ITS to individualized tutoring for any topic where student lack.

3. Conditional probabilities of auxiliary node given related concepts. Conditional probabilities are computed on the basis of importance of each concept in aggregated auxiliary node. The importance of each concept is decided by weight of the concept in associated test item with auxiliary node.
Let $\{C_{ij}, i = 1, ...n_j\}$ be the set of related concept in question associated with given auxiliary node B_j, and w_{ij} represent the importance of concept C_i in auxiliary node $B_j, i = 1, ...n_j$. Then, for each $S \subseteq \{i = 1, ...n_j\}$ required conditional probability distribution is given by:

$$P\left(B_j / \left(\{C_{ij} = 1\}_{i \in S}, \{C_{ij} = 0\}_{i \notin S}\right)\right) = \frac{\sum_{i \in S} w_{ij}}{\sum_{i=1}^{n_j} w_{ij}}$$

4. For relationships among concepts and test items, the parameters need to specify are the conditional probability of correctly answering questions given related concept and it will be fixed for every test item. That can be specified as 1 - s give related concept mastered. where s is some slip. and s for related concept not mastered.

5. For relationships among auxiliary node and test items will be same a relation between concept and test item.

4.3.3 Bayesian Adaptive Tests

As we have seen the basic for implementing Computerised Adaptive Tests(CAT) in previous section. In this section, we present Adaptive Testing based on BNs, that allows diagnosing abilities at conceptual level.

Adaptive test will be implement on network structure present in fig. 4.2. In which the knowledge nodes are concepts, topics, and subjects, and the evidential nodes are test items. The network is structure in two phases: diagnosis and evaluation phase. Diagnostic phase contains the concepts and the relationships with test items. If the test item involves with more than one concept, then the diagnosis is carried out using auxiliary nodes defined above. The goal of diagnosis is to determine the concepts that the student knows/does not know from the answers given to related test items. In evaluation phase probabilities will be propagated to determine the knowledge level at the different levels of granularity, that is, the knowledge level in each of the topics and in the subject. the adaptive test performed is content balanced so the items are selected from every concept with considering the adaptive item selection.[17]

4.3.3.1 Basic elements of the Bayesian adaptive testing

Basic elements required to implement CAT have already seen in previous section. In this section we will describe the elements to carry out Bayesian adaptive tests.

- Item Response Model-The item response model in this context is given by the conditional probabilities of the item given related knowledge node.
- Scoring Method- Scoring in Bayesian model is performed by estimating knowledge level of the concept with answers given by the student using probability propagation.
- Item Pool- Each concept must contain reasonable size of the items with all level of difficulties for more precise measurement of the ability.
- Initial Level- Initial level is set with use of available information about the student. otherwise use a uniform distribution.
- Item Selection Criteria -Adaptive criteria are used to select the best question to ask given the current estimation of the student's knowledge level. The aim of using such criteria is to achieve precise estimation of the knowledge level with shorter test item. There are several criteria are proposed, out of the the criteria based on information gain is most suitable for adaptive testing. In which next item is selected which provides maximum information gain. The Information gain defined for evidence(test item) E given knowledge node C is

  U(P,C)=(P(E=1—C=1)-P(E=1))P(C=1) + (P(E=0—C=0)-P(E=0))P(C=0)

  In above definition conditional probabilities of the evidence given knowledge node and probabilities of mastering concepts are specified in the network. And probabilities of evidence is positive is decided with the difficult of the item, which is calculated with the student responses for the given item.

  Using above specified criteria each time the question will be selected according to the performance shown by the student in the previous questions. The same criteria is used for for the evidence which depends on more than one concept. In such cases information gain is computed of the evidence E for auxiliary node B. In such cases probability propagates through auxiliary nodes to respected concept involved in question.

4.4 Alternative approach for Bayesian student modeling

In previous section we have seen application of computer adaptive test in student modeling. In that approach diagnosis is performed with adaptive testing with very large question pool. But sometime in real time environment there is not always having a possibility of large set of question for each concept to perform adaptive testing. So it is not always possible to define adaptive testing. In such cases the diagnosis is performed with available student data of access of various reading material, responses for the question from quizes and exams, assignments and exercises etc.

In certain situations the data required for accurate diagnosis of estimation of the student knowledge at every concept is not sufficient. In such situation diagnosis becomes most difficult process. This situations are well handled by managing uncertainty with Properly defining Bayesian network.

By considering the fact of knowledge propagation, the concept is mastered only after mastered all the prerequisite concepts. Using this fact Bayesian network is model with prerequisite relations. Prerequisite relationships are very useful in student modeling. [3] For example, if C_1 is a prerequisite of C_2 , and evidence shows that the student has not learnt C_1 yet, then conclusion made is C_2 is not known either. Another example can be if there are very less number of evidences of mastering C_1 with those evidence model is not certain about the knowledge of student of concept C_1, using the prerequisite relation mastering of C_2 provides the evidence of mastering the C_1. [15]

In this section we will define Bayesian Student model which will consider each small evidence that affects the certainty of the knowledge estimation. In addition to this for more precise estimation the prerequisite relations play more crucial role in short of evidences, to provide certainty in knowledge estimation. Implementation of the Bayesian network model for proposed approach is similar to the above specified approach, except the few changes in specifying conditional probabilities for prerequisite relations and conditional probabilities associated with different evidences considered in this model.

4.4.1 Nodes and relationship amongst nodes

Nodes consider for defining the model are same as defined in previous model. In this model the evidential nodes are not only the test items, they can be material access, time spend, questions answered in exams and quizes etc. There will be no change in other nodes. relationship defined in previous model will remain in this model, in addition prerequisite relation will be defined amongst concepts,between topics and concepts, and between topics. The Bayesian network is depicted in figure 4.3.

4.4.2 parameter specification

Parameter specification [15] for this model will be same as previous model. some changes and additional specification are:

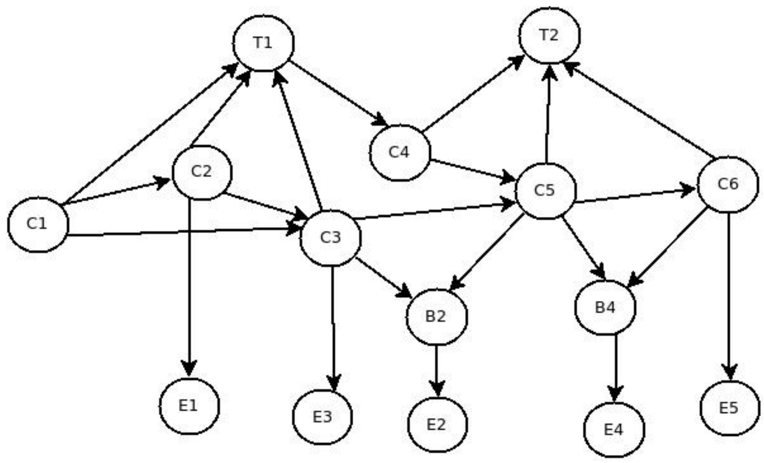

Figure 4.3: Bayesian Network model

- Prior probabilities will be assign uniformly for the concepts those don't have any prerequisite concept.
- For specifying CP of concept given prerequisite(concepts and topics) need to be specified by the expert. Machine learning techniques can be use to learn these CP, but CP calculated with these techniques are arguable.
- CP between Knowledge node/auxiliary node and evidence are assigned on the basis of how much information the evidence is providing. Evidence we are considering are Access of learning material, time spend, and questions, assignments etc. CPs for questions are assigned on the basis of the hardness level of the question. Hardness of the questions are divided in easy, medium, hard, advanced and expert scale. For each level of hardness CPs assigned with different values.

Chapter 5

System Architecture

In this section, we will see architecture of the Intelligent Tutoring System. We outline the major components of our system and describe how they interact with each other in introduction section. We discussed how to use a Bayesian network for student modelling for more accurate diagnosis of student knowledge estimation, and seen different adaptive and intelligent techniques used so far. With the study of ITS requirement, the proposed system architecture is depicted in Fig. 5.1.

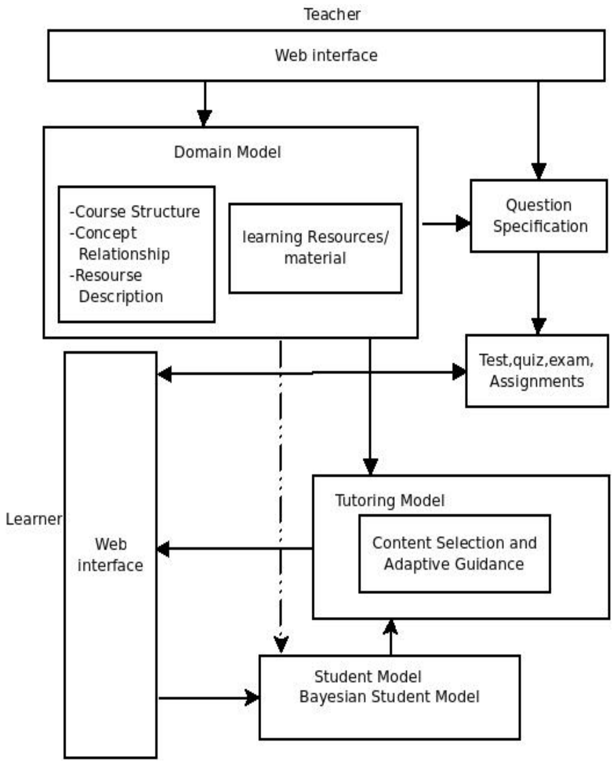

Figure 5.1: ITS Architecture

Architecture of the ITS represents different component and their working in the system. As already seen the important component of ITS in previous section, in this section we will see the working of the system with all the components and interaction with instructor and learner.

5.1 Role of Instructor and Learner

ITS presented above have assigned the roles for the Instructor and Learner are:

- *Instructor*: ITS support to offer various course on the platform. In system Instructor(teacher) will be responsible for starting the course. Instructor has to provide the concept map and learning material associated with concept, for the assessment the questions has to provided by the instructor on various concepts. As we have seen in construction of the student model, instructor has to assign the conditional probabilities among the concepts, weights of the concepts in topic, weights of the concepts in associated questions. Once all the specification provided by the instructor rest of the job will be handled by the ITS, only instructor need to initiate quizes, test, exams etc.

- *Learner*: ITS goal is to provide personalise guidance for each learner. learner study the provided learning material and with this knowledge participate in quizes, exams, assignments etc. On the basis of the learners response for the questions ITS provides Adaptive guidance to each individual.

5.2 ITS components

We have seen that ITS components include Domain model, Student model, Tutoring model and Interaction model. Here we will see how they are implemented in our system.

- *Domain model*: The content subjects to be taught are kept in the Domain Model structured as a knowledge map and a set of learning material. Knowledge map contains concept, relationship between concepts, information of associated learning material and questions etc. learning material are present in the form of lecture notes, concept pages, videos etc. This representation of model, creates separation between the domain application dependent part and the generic part. Each concept in Knowledge map correspond to the node in Bayesian network.

- *Student model*: In previous section we have seen the importance of student model in system and implementation of the Bayesian student model. For our system we will use the same student model implemented on the overlay approach in which each node represents knowledge of the student for the respective concept associated with node.

- *Interaction model*: A learner interact with system trough the web interface. In this model system collect the learner interaction data with the system. Which consist of accessed learning material information, time spend by student, responses for the questions etc. During teachers interaction with the system provide the interface to collect the domain knowledge(subject information and learning material).

- *Tutoring model*: This model is responsible for providing adaptive personalised guidance for each individual. It collects the student data from the student model and on the basis of student knowledge state it suggest the appropriate learning material from domain model. The various adaptive guidance techniques used are presented in the next section.

5.3 Adaptive Guidance

Using the state of knowledge of the student, system can offer tailored pedagogical options that support the individual student. In this section, we describe three techniques to provide personalised guidance are: navigation support, prerequisite recommendations, and adaptive presentation. [3]

5.3.1 Navigation Support

The system contains navigation menu is used for navigation through the concepts and learning material. To help the student browse through learning material, system mark each concept with some colour or signal. This signals are computed from the Bayesian network and indicate the student knowledge in that concept. In system each concept is marked in one of the category from 1)Already Known, 2)Ready to Learn and 3)Not ready to Learn.

A concept is considered already known if the Bayesian network indicates the probability p(concept = known—evidence) is greater than or equal to 0.70 or some criteria on the basis of the other student performance in the system. The value 0.7 for concept to be known is arbitrarily chosen. If a concept is marked ready to learn, it means that all prerequisites are known by the student. In Bayesian network, the probability p(concept = known—evidence) is less than 0.70 for the current concept and all of the parent concepts are already known. If at least one of the prerequisite of the concept is not known then concept will marked as not ready to learn. [3]

The direct guidance method also possible with these setting of knowledge representation. In which system will adaptively select the next most appropriate concept to learn.

5.3.2 Prerequisite Recommendations

After reading the lecture notes of a ready to learn concept, it turns out that student does not understand the concept with the assessment of the student knowledge. In such situations, system will present links to prerequisite concepts of this topic, such as the links to each concept in prerequisite in the Bayesian network. Instead of repeating the same concept again and again, this approach provides the flexibility to revisit the prerequisite concepts and confirm they are understood. Lacking prerequisites may affects the student's understanding of the concept.

5.3.3 Adaptive presentation

This technique plays an important role in adaptive personalised learning environment, which has the wide verity of learners each with different requirement and background. Technique is use full when learner satisfies all the requirement to learn the concept, that is, know all the prerequisite concepts, but unable to understand the concept even after learning specified learning material. In such cases ITS does the actions are:

- Providing explanation variants: The same information is presented in different way depends on understanding of the learner. The concept is presented in simplest form by explanation for all the concept points. The only problem with this approach is learner has to study more learning material than regular learning material for understanding of the same concept.

- Provides additional learning material and task to perform which is not part of the course but help him to understand concept. Such additional material has to be provided by the instructor and associated with concept during domain knowledge building.

Chapter 6

Conclusion and Future work

In this work we have presented modular architecture of an Intelligent Tutoring System that utilizes Bayesian networks, a proven framework for uncertainty management in Artificial Intelligence. The proposed system diagnose the knowledge of the student and provides adaptive guidance accordingly. For the development of the system we have studied different adaptive and intelligent technologies, on the basis of this study we have utilized some of the technologies to provide personalized and adaptive learning with the student state of knowledge. Diagnosis of the student state of the knowledge is most difficult task, because of the uncertain and imprecise data is managed using Bayesian network. Provided the simplified approaches for defining conditional probabilities in BN, which allows automatically computes conditional probabilities from set of weights.

Future work is to build the system which can provide Intelligent Tutoring for the MOOCs courses. In this work we have proposed the architecture of the system. In next work we are looking for the complete system which can offer personalised learning for any course in the MOOCs. We will find alternatives for simplifying the work of specifying conditional probability. Will specify the well define network probabilities by testing the system on simulated student. And the final work will be integrating with the EDX platform.

References

[1] Wikipedia. Massive open online course — wikipedia, the free encyclopedia, 2014. [Online; accessed 23-October-2014].

[2] Peter Brusilovsky. Methods and techniques of adaptive hypermedia. *User modeling and user-adapted interaction*, 6(2-3):87–129, 1996.

[3] Cory J Butz, Shan Hua, and R Brien Maguire. A web-based bayesian intelligent tutoring system for computer programming. *Web Intelligence and Agent Systems*, 4(1):77–97, 2006.

[4] Wikipedia. Intelligent tutoring system — wikipedia, the free encyclopedia, 2014. [Online; accessed 6-September-2014].

[5] Peter Brusilovsky. Adaptive hypermedia. *User modeling and user-adapted interaction*, 11(1-2):87–110, 2001.

[6] Paul De Bra, Peter Brusilovsky, and Geert-Jan Houben. Adaptive hypermedia: from systems to framework. *ACM Computing Surveys (CSUR)*, 31(4es):12, 1999.

[7] Peter Brusilovsky et al. Adaptive and intelligent technologies for web-based eduction. *KI*, 13(4):19–25, 1999.

[8] Peter Brusilovsky, Elmar Schwarz, and Gerhard Weber. Elm-art: An intelligent tutoring system on world wide web. In *Intelligent tutoring systems*, pages 261–269. Springer, 1996.

[9] Peter Brusilovsky and Christoph Peylo. Adaptive and intelligent web-based educational systems. *International Journal of Artificial Intelligence in Education*, 13(2):159–172, 2003.

[10] Peter Brusilovsky and Eva Millán. User models for adaptive hypermedia and adaptive educational systems. In *The adaptive web*, pages 3–53. Springer-Verlag, 2007.

[11] Constantino Martins, Luíz Faria, Carlos Vaz De Carvalho, and Eurico Carrapatoso. User modeling in adaptive hypermedia educational systems. *Educational Technology & Society*, 11(1):194–207, 2008.

[12] Loc Nguyen and Phung Do. Learner model in adaptive learning. *World Academy of Science, Engineering and Technology*, 45:395–400, 2008.

[13] Hugo Gamboa and Ana Fred. Designing intelligent tutoring systems: a bayesian approach. *Enterprise Information Systems III. Edited by J. Filipe, B. Sharp, and P. Miranda. Springer Verlag: New York*, pages 146–152, 2002.

[14] Eva Millán, Mónica Trella, José-Luis Pérez-de-la Cruz, and Ricardo Conejo. Using bayesian networks in computerized adaptive tests. In *Computers and Education in the 21st Century*, pages 217–228. Springer, 2000.

[15] Eva Millán, Tomasz Loboda, and Jose Luis Pérez-de-la Cruz. Bayesian networks for student model engineering. *Computers & Education*, 55(4):1663–1683, 2010.

[16] Ricardo Conejo, Eduardo Guzmán, Eva Millán, Mónica Trella, José Luis Pérez-De-La-Cruz, and Antonia Ríos. Siette: A web-based tool for adaptive testing. *International Journal of Artificial Intelligence in Education*, 14(1):29–61, 2004.

[17] Eva Millán and José Luis Pérez-De-La-Cruz. A bayesian diagnostic algorithm for student modeling and its evaluation. *User Modeling and User-Adapted Interaction*, 12(2-3):281–330, 2002.