Profile Based Electronic Learning Content Aggregation and Delivery System Using Self Learning Satisfaction Index of User as Feedback

Nitin Pujari, Vinay S K, Sucharitha Prabhakar

PES Institute of Technology, India

0403

The Asian Conference on Technology in the Classroom 2013

Official Conference Proceedings 2013

Abstract

Every human entity born in the universe, will possess distinct characteristics, some inherited, other their own. Human beings when compared mainly differ in their biological, chemical and psychological characteristics. These differences make human beings to receive, perceive, interpret and store information differently based on their cognitive abilities. This proposes a challenge for the design of traditional and/or electronic teaching learning systems to cater to the diversified needs of the learner. In traditional learning systems, the challenge is on the tutors to identify learning abilities of the pupils and teach accordingly. To some extent this has been achieved by good tutors, which most of the time is not a scalable proposition. The advent of electronic learning systems has been able to successfully achieve the scalability aspect of this ecosystem. The current electronic learning systems to greater extent do not take into account the learning ability or comprehension index of the learner. This work attempts to propose a concept, model and/or mechanism for electronic learning systems to aggregate and deliver the content to the learner based on the profile of the user which include comprehension index of the learner. This is achieved by utilizing self learning satisfaction index of the learner as feedback.

Keywords: *social networks, context, congruence, congruent network, homogeneous, heterogeneous*

I.   Introduction

Knowledge resources generally are not organized based on the cognitive abilities of an individual. Individuals with lesser cognitive abilities will find it difficult to understand the contents presented in the knowledge resources which have higher comprehensibility. On the other hand, people with a higher cognitive ability will not find a knowledge resource with lesser comprehensibility useful. The content presented in the knowledge resources are expressed from the perspective of conveying the concept assuming certain pre-requisites which sometimes masks the cognitive abilities of the learner. It is obvious that the learners have different cognitive abilities and therefore there is a need to identify learning abilities of the learner and present the content in which he/she would feel congruent or appropriate. Learning abilities include the ability to understand, familiarity with the concept and prerequisite knowledge. To make learning process productive, this work proposes a concept, model and/or mechanism called "Profile based electronic learning content aggregation and delivery system using self learning satisfaction index of user as feedback" which takes into account the learning abilities and presents the appropriate content. The proposed concept and model can be implemented in three steps:

 1) Self-Assessment

 2) Knowledge Progression

3) Assimilate New Concepts

1) Self assessment: Learner is expected to have some understanding of the domain/sub-domain during self assessment. The proposed system encourages different approaches for automating the generation of a question bank based on the ability of the user. This can be achieved by analyzing the quality of basic knowledge that he/she posses represented by a paragraph. The learner evaluates and gauges the level of knowledge and gives feedback to the system for the next step of actions.

2) Knowledge Progression: Learner assesses his conceptual knowledge from step 1. Based on the assessment, learner can deduce the self learning satisfaction index. If self learning satisfaction index is high, learner can concentrate on a different concept, or else can decide to increase his expertise in the chosen concept after comparing with the performance index generated by the system.

3) Assimilate New Concepts: The learner can assimilate new concepts in a chosen domain. In this case, learner selects the concept in which learner is a naive. The proposed system treats the learner as a beginner. Initially an introduction to the concept is presented. The learner attempts to understand the concept. On success learner can proceed with knowledge progression or else the system will suggest pre-requisite concepts. Learner can proceed with the knowledge progression or assimilation of the pre-requisite concepts.

The system we are proposing uses profile as the basis for content selection and assessment. All the result of activities with our system such as feedback will be recorded in the profile.

Profile of a user can be defined as the information of the user which will help us in fine tuning the content for learning. Profile can also be treated as a tool for personalized learning for each user. In our proposed system, the profile of a user is

not constant / not fixed. For every learning process, a profile is necessary. The profile must consist of a specification of information related to the domain and subject.

The proposed profile based system differs from the other existing systems, in such a way that our system takes into consideration, multiple profiles for each user one profile for each learning of new domain.

II.   Architecture and Overview of the Proposed System

Sources for content aggregation and delivery are collected and organized in a container called as the Universal Knowledge Base (UKB). The UKB has all the resources that are required to present the appropriate content. These resources include the following:

- Textbooks
- Links for portals and discussion forums
- Research articles
- Presentations
- Universal knowledge classifications

All of these resources are considered as the "Knowledge resources". Each of the knowledge resource is used for a specific purpose at different stages. For example, links for portals and discussion forums are used to generate questions for self-assessment. The proposed system can be implemented by logically sequencing the following steps:

1) Knowledge Base Organization
2) Weighted Score Assignment
3) Determining Learner Comprehension Index and Content aggregation
4) Question bank generation
5) Self-Assessment, Feedback Processing and Content delivery

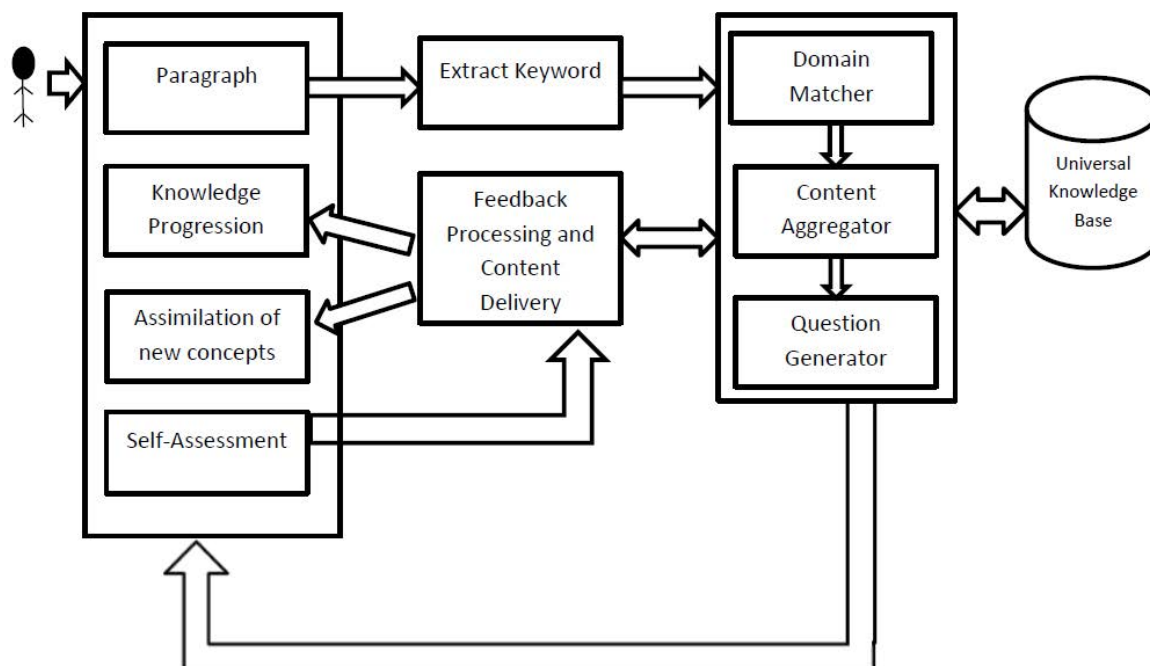The architecture of the proposed system is as follows.

Figure 1: Architecture of the proposed system

Upon the organisation of the universal knowledge base, each of the knowledge resource will be assigned a score in the weighted score assignment step. This score helps to filter out the contents that are appropriate for the learner. In the third step, where the learner comes into picture for the first time, analysis of the learner is made to find the best available contents for the learner. It is done on the basis of the comprehension index of the learner. Learner is asked to express knowledge of the content which learner wishes to learn, in textual form. From the paragraph, comprehension index of the learner is determined. Based on this analysis, content from the knowledge resources are aggregated. From the aggregated content, a set of questions are generated to assess the level of understanding of the user. The proposed system differs from existing systems in a way that, assessment is made by the learner. This helps to determine the "self learning satisfaction index" of the learner. Based on this index, used as feedback, further assessment and content delivery is made.

The usage of profile is done in 2 stages-

1. *Profile creation:* In this stage, the user is going to enter details such as achieved proficiency in domains, subjects, in which the user is interested. These fields will be treated as static.

2. *Profile Evolution*: In this stage, evaluation is done and the profile is updated. For each learning process, the following fields will be added to the profile –

    Questions: Number of questions generated for the user entered paragraph.

Performance Index: Number of questions that are answered by the user. This is the field where   feedback of self assessment is recorded.

Recommendation: Recommendation of the number of questions to be answered by the user, based on the profile created in the first stage.

In addition to the assessment through question-answering process, the user can select the required proficiency in the concept. The selection of proficiency should be made before the self assessment stage. The value of the 'recommended' field mentioned above changes with the selection of proficiency.

After each assessment, feedback of the user is recorded in their profile for further learning process.

In the 'Assimilation of new concepts' phase, the pre-requisite courses will be recommended by the proposed system. Initially, the Achieved Pre-requisite Index (API) will be calculated after analyzing the subjects or domains from the initial profile. The calculated API is further refined as the profile evolves with learning process.

III.   Knowledge Base Organization

Universal Knowledge Base has classified set of domains. This keyword based classification is done using a set of qualifiers. Each domain encompasses a set of subjects. Subjects are assigned to a domain based on the keywords it represents as a subset of domain based keywords. Each subject has a particular set of knowledge resources. Each knowledge resource has a set of keywords representing the knowledge. Further, Knowledge resources are organized in terms of concepts, each one covering a part of the knowledge. Concepts are presented as set of multiple paragraphs; each paragraph represents a portion of a particular concept. The entire universal knowledge base can be organized hierarchically as shown below.
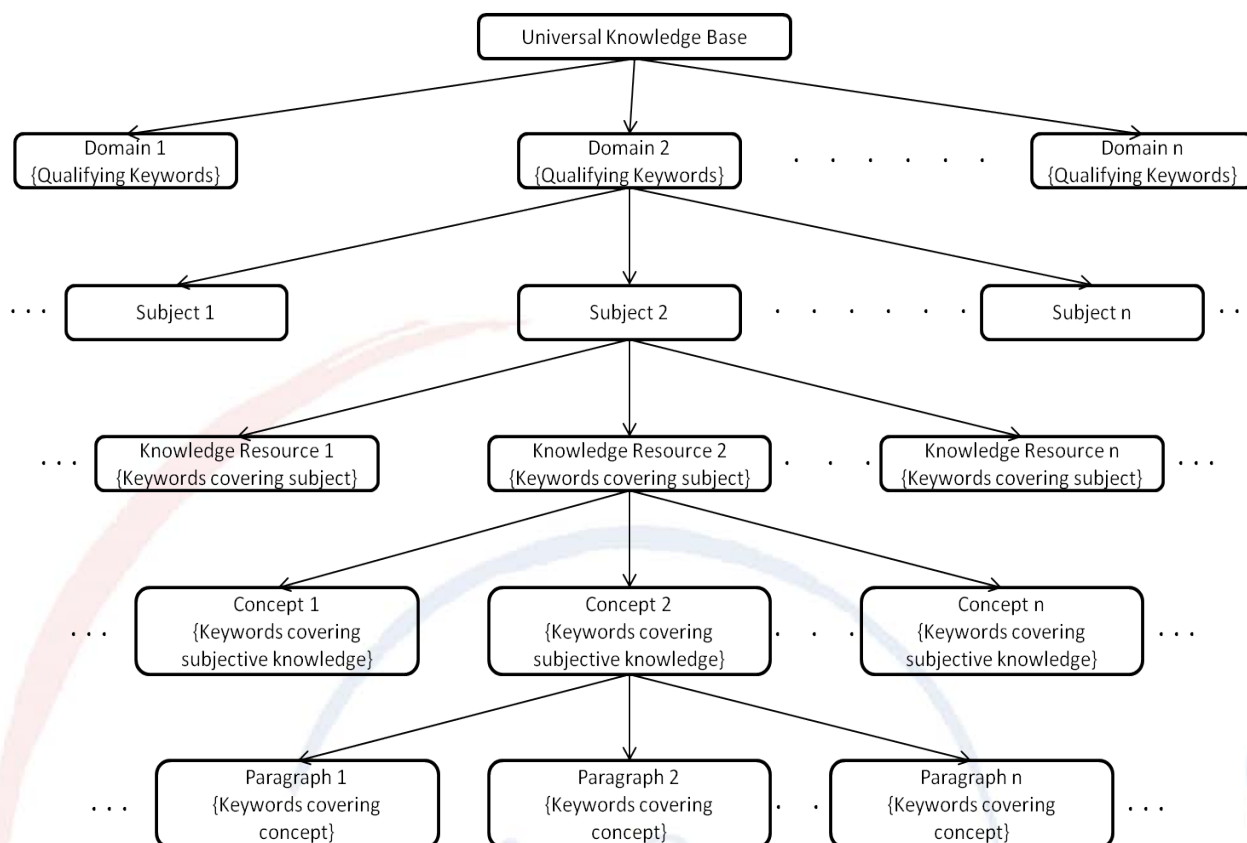
Figure 2: Knowledge Base Organization

From the above figure, we can list the components of knowledge base as follows

1. Domain
2. Subject
3. Knowledge Resource
4. Concept
5. Paragraph

Domain: Domains can be obtained from the available knowledge sources such as Wikipedia, DBPedia etc.   We are using the available classification from the sources. Example for a domain could be "Computer Science".

Subject: Subject covers a portion of the domain. Example subjects that are under the computer science domain could be data structure, operating systems etc. Subjects may span across multiple domains. For example, microcontrollers come under both computer science and electronics domain.

Knowledge Resource: Subjects are described in various knowledge resources. Knowledge resources are the source of content for learning.

Concept: Each knowledge resource presents the content of the subject in-terms of concepts. Concepts are similar to chapters in the text-book.

Paragraph: Concepts are expressed in different paragraphs. Each paragraph is a distinct section in a concept.

IV. Weighted Score Assignment

The proposed system uses weighted scores to choose contents for aggregation and delivery. When the universal knowledge base is organised, scores are assigned. Assignment of scores is done based on the coverage factor which can be defined as the relative amount of knowledge a particular item covers with respect to a domain it belongs. From the coverage factor, scores at each level can be assigned as follows.

Subject Score = *keywords covering domain / Total keywords in the subject*

Knowledge Resource Score = *keywords covering subject / Total keywords in the knowledge resource*

Concept Score = *keywords covering knowledge resource / Total keywords in the concepts*

Paragraph Score = *keywords covering concept / Total keywords in the paragraph.*

In addition to assigning scores based on coverage, knowledge resources are further classified to distinguish between the "levels of specialization"/ "the level of difficulty". The classification is as follows

1. Text-books – for regular reading
2. Reference books – for obtaining deeper knowledge with much elaborated contents
3. Research Oriented books – for highly specialized reading

To distinguish the specialization level, knowledge resources are assigned weights. 1 for text-books, 2 for reference books, 3 for research oriented books. These weights are called 'specialization weights'. The scores are stored in the knowledge base itself.

V. Determining Learner Comprehension Index and Content Aggregation

The learner demonstrates knowledge of the concept by entering text related to the concept. The model extracts all keywords from the text. Based on the extracted keywords, the domain to which the concept belongs to can be inferred by referring to the universal knowledge base. A Score is assigned for the paragraph. These scores together are used to determine the index called "Comprehension Index" of the user. The Comprehension Index is used to determine the proficiency of the learner in the context of the concept. The scores are assigned based on two factors called "strength" and "weight". Strength measures the breadth wise knowledge and weight measures the depth wise knowledge. The whole process of learner analysis and content aggregation proceeds in the following stages.

1. Domain Filtering
2. Subject Filtering
3. Knowledge Resource Filtering
4. Concept Filtering
5. Paragraphs Filtering

*Domain Filtering***:** The aim of domain filtering is to select a domain where the learner written concept belongs.   It is done with the help of strength and weight factors.

Domain Strength (DS) = *no. of domains covered/total domains in the universal knowledge base*

Lesser the value of domain strength implies that the textual content given by the learner falls in a domain. Larger the value, higher is the number of domains covered. In that case, the model asks the learner to re-write the concept. The process is repeated until the domain strength asserts single domain coverage. The above mentioned process is not expected to repeat because the entered content may not span across multiple domains.

Then the Domain Weight is calculated for each domain as follows

Domain Weight (DW) = *no. of keywords covering the domain / total keywords in the content of learner.*

The model selects the domain which has the highest value. It is guaranteed that, only one domain will have largest weight because the domain strength is less.

Subject Filtering: Once the domain has been selected, the model repeats the strength and weight calculation for the subject.

Subject Strength (SS) = *no. of subjects covered/total subjects in the domain*

Lower subject strength implies that the content expressed by the learner falls in a subject. Larger the value, higher is the number of subjects covered. In that case, the model asks the learner to re-write his concept. The process is repeated until the subject strength asserts single subject coverage. Similar to domain coverage, the process is not expected to repeat.

Similar to domain weight, the model calculates subject weight for each subject and the subject with highest subject weight is selected.

Subject Weight (SW) = *no. of keywords covering the subject / total keywords in the content of learner*

Similarly, the knowledge resource filtering and concept filtering are done using the following formulae.
Resource Strength (RS) = *no. of knowledge resources covered / Total no. of knowledge resources in the subject.*

Resource Weight (RW) = *no. of keywords covering the knowledge resource / total keywords in the content of learner.*

Concept Strength (CS) = *no. of concepts covered / Total no. of concepts in the knowledge resource.*
Concept Weight (CW) = *no. of keywords covering the concept / total keywords in the content of learner.*

After the above steps, the model would have deduced the concept expressed by the learner. As shown in the knowledge base organization, a concept can consist of multiple paragraphs. For each paragraph, paragraph weight is calculated using the formula

Paragraph Weight (PW) = *no. of keywords covering the paragraph / total keywords in the content of learner*

For paragraph filtering, the model uses a threshold value which is the average paragraph weight.
Threshold = $\sum PW$ / *no. of paragraphs*.
The proposed system, considers paragraphs whose PWs are above the threshold.
The content that is considered for question generation and later delivery is an aggregation of paragraphs selected in the previous stage.

VI.   Question Generation

Once the domain and the subject are inferred, the model fetches all questions related to the concept which the learner wishes to learn, from the question resources which are associated with the subject. The question resources include question and answer portals and discussion forums. To generate questions, we are using the internally aggregated content. From this content, all keywords are extracted. Questions will be fetched from all the available question resources .The obtained questions for the user is an aggregation of the questions containing the keywords from all question resources.

A)   Eliminate Redundant Questions

It is very likely that the same question will exist in multiple question resources. Therefore it is necessary to filter the questions obtained from each question resource. The model follows a two step filtering process.

Type 1 Filtering: Questions which contain the same words are redundant. All questions which have the same form, that is contain the same words are eliminated. Questions which do not contain the same words may also have the same meaning. Therefore redundancy still exists in the obtained question set. To eliminate this, the model goes for the second step of filtering.

Type 2 Filtering: Questions which are semantically equivalent to each other are redundant. The model eliminates all questions which are semantically equivalent. At the end of this two step filtering process we obtain   the question set which contains questions which are distinct from each other.

VII.   Self Assessment, Feedback Processing and Content Delivery

The learner is presented with the questions which were generated in the question generation phase. The learner attempts to answer the questions. The learner evaluates himself by calculating the user self satisfaction index.

User self satisfaction index=*number of questions answered correctly/total number of questions presented.*

The user self satisfaction index of the user for this concept is archived by the model. The learner has to determine if he is satisfied or not satisfied. This decision is a subjective measure and varies between users. The self satisfaction index of the user is taken as feedback. Feedback processing is performed as follows.

A)   *If the learner is satisfied :*

The learner can switch to a new concept and start with either knowledge progression phase or assimilation of new concepts.

i) If the learner wishes to start the knowledge progression phase of a different concept the learner has to express his knowledge of the concept in textual form, answer the questions presented to him, calculate the user self satisfaction index and proceed with the knowledge progression phase if he is not satisfied or switch to assimilation of new concept.

ii) If the learner wishes to start the assimilation of new concept, he has to specify the concept. The model determines the prerequisite knowledge required. This is determined by traversal of the global knowledge base. The traversal starts from the root node in the tree till the required concept node is found. All siblings of this concept node which appear to the left of it and are at the same level are the required prerequisite concepts. The model evaluates the achieved prerequisite index(API) of the user for this concept .The achieved prerequisite index(API) of the user is the average of the user self satisfaction index of each of the prerequisite concepts

API = *sum (user self satisfaction index of prerequisite concept) / count (prerequisite concept)*

An API of 5 is considered the minimum score required to proceed with the assimilation of the new concept.

 If the API is below 5 the learner is expected to proceed with the assimilation of the prerequisite concepts if the user is not familiar with the prerequisite concept or the learner can proceed with the knowledge progression phase of the perquisite concepts.

The learner can proceed with either of the above phases till the API is greater than 5 for the learner. Once the API is above 5, the learner can proceed with the assimilation of the concept which he desired to learn initially.

B)   *If the learner is not satisfied:*

That is user self satisfaction index is low, the user is taken into the knowledge progression phase and presented with the content which was aggregated for the user.

The user is expected to learn the concepts presented to him and then express his understanding of the concept in a textual form. The user entered source is evaluated and the user comprehension index is evaluated. The user comprehension index of the user is expected to increase.

a) If the user comprehension index for this concept has not changed the user is presented with the same concept till the user comprehension index of the user increases.

b) If the user comprehension index has increased, questions which are generated for the advanced concept are presented to the user. Self-assessment of the user takes place. The user evaluates his user self satisfaction index and based on it, he can proceed with the knowledge progression or assimilation of new concept.

References

[1] Pei-Chen Sun, Ray J. Tsai, Glenn Finger, Yueh-Yang Chen,Dowming Yeh, "What drives a successful e-Learning? An empirical investigationof the critical factors influencing learner", Science Direct, doi:10.1016/j.compedu.2006.11.007

[2] y. matsuo, m. ishizuka, "keyword extraction from a single document using word co-occurrence statistical information", International Journal on Arti_cial Intelligence Tools, Vol. 13, No. 1 (2004) 157-169

[3] Frank, E., Paytner, E., Witten, I.H., Gutwin, C, and Nevill-Manning, C.G. (1999). "Domain specific keyphrase extraction". In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence, Morgan Kaufmann, pp:668-673.

[4] Turney, P. (2000). "Learning to extract keyphrases from text". Journal of Information Retrieval 2(4):303-336.

[5] Witten, I.H., Paynter, G.W., Frank, E., Gutwin, C, and Nevill-Manning, C.G. (1999). KEA: "Practical automatic keyphrase extraction" In Proceedings of DL '99, pp:254-256.