ISSN: 1813-4890

Personalized Information Retrieval to Scaffold Learning

Qianyi Gu
School of Computer Science, Sichuan Normal University, Chengdu 610101, China.
qianyigu@foxmail.com

Abstract

With the rapid development of online education, many learners go to internet to find materials they need. Considering the vast amount of digital educational resources available, the learning challenge becomes to go beyond the ability to locate information and to acquire meaning and gain understanding. The paper explores how to design an intelligent learning service based on the learners' prior knowledge that would scaffold the learners effectively in improving their scientific understanding. The paper proposes scaffolding strategies for education developers to help learners search and effectively use online educational resources. To support these strategies, personalized information retrieval are implemented. The algorithm is based on what learners already know and what they should know in their learning process. The artificial intelligence techniques are used to construct the knowledge model based on concept matrix. The learning activities are modeled to automatically detect the learner's own learning needs. The intelligent learning service is able to support a deeper integration of science content with the learner's prior knowledge.

Keywords

Information retrieval, online learning, scaffold strategy, knowledge articulation.

1. Introduction

Along with the arrival of the internationalization age, e-learning based on digital educational resources has received widespread attention. A large number of resources are built by teachers and publishers[1]. One of the key challenging tasks for educators is to provide learning services to support learners to effectively locate and use the available educational resources to fulfill their learning needs and improve their understanding[2]. Learners have varying states of background knowledge, learning preferences, and educational levels. Research in cognitive science found that differences in individual learners' knowledge backgrounds and learning styles affect their learning processes and ultimately their effectiveness[3]. Learners enter each learning topic with their own preconceptions about the knowledge to be learned and its related concepts[4]. If the educators can develop teaching strategies and design for each learner's preconceptions and learning needs, best learning results are expected to be achieved[5].

With the vast amount of learning resources available online, key words are not always an effective method of locating the resource to fulfill the learning needs. Researchers have shown that information retrieval will be more effective if the individual learner's learning requirements are taken into account[6]. Learners build their new scientific understanding based on their preconceptions and affected by their previous misunderstanding. Taking prior knowledge into account is a critical factor for success in learning. This paper proposes a personalized learning service based on learners' prior knowledge and misconceptions. To achieve personalized information retrieval in online learning environments, understanding the learner's needs becomes a central question upon which to build such retrieval service. Furthermore,when the learner is placed at the center, the goal of information retrieval is no

ISSN: 1813-4890

longer merely getting the relevant information source but getting information that leads to a new understanding in the process of learning[7]. The goal of the paper is to design and implement a personalized learning service to scaffold learners effectively using online educational resources to enable the intentional learning process.

2. Scaffold Strategies

The goal of the information search is to collect information that leads to a new understanding in the e-learning process. Thus, selecting resources by purely forming a set of search terms and passing it onto the search engine cannot fulfill the requirements. The information retrieval techniques need to achieve three requirements. They need to be able to take into account an individual learner's prior knowledge as a basis of forming strategies to find information resources to scaffold the learning process. They need to be able to identify the relationships between suggesting educational resources and an individual learner's current concepts and knowledge. They need to be able to form a set of strategies to find resources to correct an individual learner knowledge gap and misunderstanding of certain concepts. Two scaffold strategies are proposed in the paper to achieve these requirements.

2.1. Boundary Strategy

Prior research has shown that learners are frequently in the situations in which they are overwhelmed by the huge number of online educational resources available. In such situations, they may have difficulties starting the search process to locate the right information and build new understanding to complete their tasks. The boundary scaffold strategy is aimed at helping learners decide whether the resources they have gathered are enough for them to achieve their current learning tasks. It helps them to form a focused perspective upon which to build and to decide what information pieces are to be included in their learning process for the new scientific understanding.

The boundary strategy suggests that educational resources to be retrieved based on a particular concept or certain concepts that are related to each other. Linking the retrieved resources to a group of concepts will help learners explore a general topic for an aspect to pursue in the information seeking process and to form a focused perspective.

2.2. Extension Strategy

When learners are investigating various online educational resources, they usually find different resources seem inconsistent and incompatible and they feel uncertain and confused about this inconsistency. The extension scaffold strategy is proposed which provides support to the learners by relating the new information to their existing understanding to conquer these uncertainties.

The extension strategy suggests that educational resources to be retrieved based on an individual learner's preconceptions and knowledge gaps. Identifying the relationship between retrieved resources and individual learner's preconceptions will facilitate the learners in relating the new information to their existing understanding. While identifying the relationship between retrieved resources and the individual learner's knowledge gaps will facilitate the learner's understanding of how the new information can be used to inform their existing misunderstanding.

3. Conceptual Framework of Personalized Information Retrieval

The paper proposes a conceptual framework of the personalized information retrieval service based on learner's prior knowledge. There are three core dimensions of this framework: knowledge model, resource model and knowledge need articulation.

3.1. Knowledge Model

The knowledge model describes how the domain and learner knowledge are represented in the personalized information retrieval service. A conceptual model and a computational model are used to represent the domain knowledge in the service. The conceptual model is a knowledge map representation of the domain and learner knowledge, including the learner's misconceptions. The computational model is a concept matrix transformed from the knowledge map representation. The learner concept graph, which represents the learner's knowledge status, prior misconceptions, and learning needs, is constructed from the concept matrix.

3.2. Resource Model

The resource model describes how the online educational resources are represented in the personalized information retrieval service. The online resources have a conceptual representation, which consists of web page content, metadata description, and a linked web page from the online resource. This conceptual representation is transformed into a computational representation in the service when the content and metadata of the resources are crawled and indexed into the service. From that index, the web page concept graph is dynamically constructed, aligned to a particular learner concept graph when personalized information retrieval is executed.

3.3. Knowledge Need Articulation

The learners' knowledge need articulation process is based on the type of the learner's knowledge gap and misconception. There are three types of learner's misconception addressed in this framework: incomplete understanding, fragmented knowledge and incorrect statement.

3.3.1. Incomplete Understanding

For the misconception of incomplete understanding, it is necessary to identify the core concepts that are missing from the learner's current knowledge as well as the important concepts related to the missing cores. The missing node in the learner knowledge map are called core knowledge node which should have a corresponding knowledge node in the domain knowledge model. Knowledge nodes in the learner knowledge map that are closely related to the core node are called learner associated nodes. Domain associated nodes are the knowledge nodes in the domain knowledge model that are related to the core node.

The concept matrix is constructed from the identified concept nodes and terms. For each key term extracted from the core node, the algorithm computes the association value between this term and terms extracted from learner associated nodes. For each key term extracted from the learner core node, the algorithm computes the association value between this term and terms extracted from domain associated nodes. For each key term extracted from the learner associated nodes, the algorithm computes the association value between this term and terms extracted from domain associated nodes.

The educational standards are used to expand the learners' knowledge. The service finds educational standards aligned to each involved node and extracts key terms from educational standards. It then computes association values of the extracted key terms with core terms.

3.3.2. Incorrect Statement

For the misconception of incorrect statement, it is necessary to identify the core problematic concepts from the learner's current knowledge as well as the important concepts related to this

ISSN: 1813-4890

core problematic concept. The algorithm identifies nodes directly connected to the incorrect node and nodes indirectly connected to the incorrect node. It then partitions terms as core terms, faceted domain terms, learner matching terms, and learner related terms.

Faceted information gathering techniques are used to expand the core domain concept identified. The algorithm gathers knowledge nodes from different directions of the domain knowledge model. For each direction with a set of nodes, it defines the set of nodes as a facet. For each facet, the algorithm picks key terms from the set of nodes in this facet as faceted domain terms.

3.3.3. Fragmented Knowledge

For the misconception of fragmented knowledge, it is necessary to identify the core disconnected concept nodes from the learner's knowledge map as well as the important concepts and terms related to these disconnected nodes.

The algorithm identifies the pair of nodes in the domain knowledge model that are missing a relationship in the learner's knowledge map. It then identifies nodes on the path between core nodes in the domain knowledge model that are aligned to nodes in the learner's knowledge map. Finally, nodes on the path between core nodes in the domain knowledge model that are not aligned to any node in the learner's knowledge map are picked.

A concept matrix is constructed from the identified concept nodes and terms. For each core term, the algorithm computes the association value between this term and related learner terms. The association value between core terms and their related unaligned domain terms are also computed.

4. Technical Implementation

The technical implementation of the personalized information retrieval consists of four major components: knowledge model processor, search strategies generator, educational resource processor and matrix analyzer.

4.1. Knowledge Model Processor

The knowledge model processor is used to process knowledge data from other applications and generate a concept matrix. The input data includes domain knowledge model, learner's knowledge maps and the learner's knowledge gap objects. The learner's knowledge gap objects contain knowledge gap type, involved learner's concepts, and involved domain concepts. The knowledge model processor takes the input knowledge data and generates a concept matrix. The concept matrix is the computational knowledge model that is used to represent the conceptual knowledge map model for the domain and the learner. It is used to provide the data structure of the conceptual knowledge model for personalized information retrieval based on the individual learner's prior knowledge. The concept matrix representations are composed of key concepts extracted from the domain knowledge model and learner knowledge model. The two dimensions of the concept matrix are both key concepts, and the value in the matrix represents the distance between two concepts. The algorithm first identifies key terms from each knowledge node in the knowledge map and assigns weights to each identified key terms. Then it computes the distance value of each pair of the identified key terms. The service merges the identical key terms to construct the concept matrix.

4.2. Search Strategies Generator

The search strategies generator is used to form a set of processing strategies based on the current information seeking task. The two types of information seeking tasks for online learning are called concept extension and misconception correction.

The goal of concept extension task is to select online sources based on the learner's missing concepts and domain concepts that are associated with the learner's existing understanding. This usually happens in the early stages of the information search process. The learner explores the information space to find a focus of topics for her information needs. In such task, search terms might be provided by the learner as an indication of topics she tries to explore. After the search strategy generator processes the learning needs, it will generate the strategies that are composed from analyzing requests of elements in the concept matrix from the concept matrix analyzer.

For the misconception correction task, a methodology to identify resource selection strategies based on knowledge gap types and expert heuristics is designed. The search strategies generator is the computational implementation of the methodology. The result of such implementation is an association of the learner's knowledge gap matrix with elements in the learner and domain concept matrix.

4.3. Educational Resource Processor

The educational resource processor is the key component to index and analyze the online educational resources. The goal of the educational resource processor is to establish the relationship between the educational resources and the learner's knowledge. The process results are used to decide whether the educational resources are useful for this learner's particular learning task and how the services should present the resources to the learners.

For each online educational resource, the processor stores the index and description of pedagogical related objects contained in the resource. It processes the stored index and generates an association of that educational resource with domain concepts and learner concepts. The final result of the process is a resource matrix that represents the relationship of concepts and that resource.

4.4. Matrix Analyzer

The Matrix analyzer is the key component to decide what educational resource to be recommended to the learner and how to present that resource to the learner. The matrix structure of the knowledge model is used to represent the quantitative relationship and patterns between different concepts in the knowledge maps, and the relationships are used in the resource selection implementation. In finding resources based on a learner's particular misconceptions, the matrix analyzer will compute the association value of learner knowledge gap matrix and resource matrix. The association value is an indicator of how the information resource relates to concepts involved in that misconception. Then instructional strategies will be used to decide whether the service will recommend that resource to the learner. In concept extension, the matrix analyzer computes the association value of the learners' concept matrix and the resource matrix. The result is an indicator of how that information resource is related to the learner's current knowledge and concepts.

5. Conclusion

The paper proposes an implementation of the personalized information retrieval service based on an individual learner's prior knowledge. The implementations are used as computational service to support the scaffolding strategies identified by educators. The description starts with the overall framework of the learning service. It consists of knowledge maps and misconceptions represented as a knowledge model, an online resources graph represented as resource model, and learner's knowledge need articulation based on the models. Then the description of the workflow of the service is provided. The key data structure used in this implementation, the concept matrix, is discussed. The methodology to transform knowledge maps into such a matrix is explained. The construction of the learner and resource concept

ISSN: 1813-4890

graphs is discussed. Finally, the technical implementation based on the concept graph is then shown. It is expected that the use of the graph based algorithm to make personalized recommendations is effective for addressing learner's individual learning needs.

References

[1] D. Chaitali, S. Srinath, S. Gandharv, et al: AI-based learning content generation and learning pathway augmentation to increase learner engagement, Computers and Education: Artificial Intelligence, 2023, Vol. 4(2), 100-110.

[2] W. Kim, L.S. Watson: Perceived learning in three MOOCs targeting attitudinal change, Educational Media International, 2016, Vol. 53(3), 168-183.

[3] J.D. Bransford, A.L. Brown, R.R Cocking: *How People Learn: Brain, Mind, Experience, and School.* (National Academy Press, Washington D.C., 2000).

[4] R.C. Clark, R.E. Mayer: *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning* (Wiley, USA, 2016).

[5] D.H. Jonassen, B.L. Grabowski: *Handbook of Individual Differences, Learning and Instruction*. (Taylor and Francis, London, 2012).

[6] C.C. Kuhlthau: *Seeking Meaning: A Process Approach to Library and Information Services.* (Bloomsbury Publishing, USA, 2004).

[7] H.N. Agreda, S.D.M. González, R.R. López: Development of Transdisciplinary and Complex Learning in Inclusive Educational Practices, Education Sciences,2024, Vol. 14(3), 222.