Scientific
Research
Publishing

Learning Path Planning Based on Knowledge Graph on MOOC Platform

Boya Zhang¹, Huimin Si²

¹Qianyuan 2nd Phase Residential Quarter, Shijiazhuang, China
²Royal Yili Residential Quarter A, Cangzhou, China
Email: tjnuzhangboya@163.com, huiminsi@163.com

Abstract

Based on the edX platform, this paper designs the knowledge graph model by combining expert method and reference method, selects Neo4j as the knowledge storage database of knowledge graph; selects Echarts as the visualization tool of Neo4j, and finally realizes the visual search and display of knowledge points on MOOC platform, and gives learning feedback and update according to the learning situation of learners. When planning the learning path through the designed knowledge graph, Ripplenet algorithm is introduced to enrich the diversity of courses in the learning path planning between courses. Professional learning path planning is provided for learners to improve the learning efficiency of online learning.

Keywords

Knowledge Graph, Learning Path, Neo4j, Visualization, Open edX

1. Introduction

MOOC platform provides strong support for learners to achieve autonomous learning and lifelong learning. Especially during this epidemic, MOOC provided convenience for teachers and students who study at home. However, the online learning platform still has various drawbacks. 1) The lack of a complete knowledge system of numerous teaching resources will cause learners to lack personalized guidance in learning (Cao & Zuo, 2018; Yang & Zhuo, 2017). 2) Lack of feedback on learners' learning effect (Zhang, 2019). To sum up, the massive learning resources on MOOC platform provide the guarantee of course resources for learners' online learning, but there are still some problems, such as the lack of management of learning resources, the inconsistency between the knowledge structure of resources and the real network knowledge structure, the lack of guidance for learners' learning, and the lack of feedback on learning effects.

Moreover, each learner's learning starting point and goal are not the same (Li, Shao, Wang, Zhao, & Guo, 2018). Therefore, it is necessary to provide learners with a learning path planning that matches their knowledge mastery and aims at achieving their learning goals. Learning path planning on MOOC platform is an important research topic of this paper.

2. Knowledge Graph Design

2.1. Semantic Type Design of Knowledge Graph

By referring to the semantic type diagram of unified medical language system and combining with the research content of this paper, the semantic types of knowledge graph in this paper are constructed.

The semantic types of unified medical language system believe that everything can be divided into things, and things can be divided into "Entity" and "Events" (APA Divisions, n.d.). Therefore, "Entity" and "Events" constitute the first level category of semantic types in the unified medical language system. According to the superior and inferior relationship between entities and concepts, and based on the first level category of semantic types, gradually downward.

The semantic type of "Entity" includes "Conceptual Object" and "Physical Object". Descriptions of some abstract things, such as educators' educational thoughts and learners' learning abilities, are placed in semantic type under the category of "Conceptual Object". The description of some things that exist objectively on MOOC platform, such as course information and course content, are placed in the semantic type under the category of "Physical Object". "Conceptual Object" and "Physical Object" constitute the main part of semantic types of unified medical language system.

The semantic type of "Event" includes "Activity" and "Phenomenon or Process". The description of the behavior and activities of educators, learners, developers on MOOC platform are placed in the semantic type under the category of "Activities"; The phenomena or descriptions of the process of interaction between educators, learners and developers on MOOC platform are placed in the semantic category of "Phenomenon or Process".

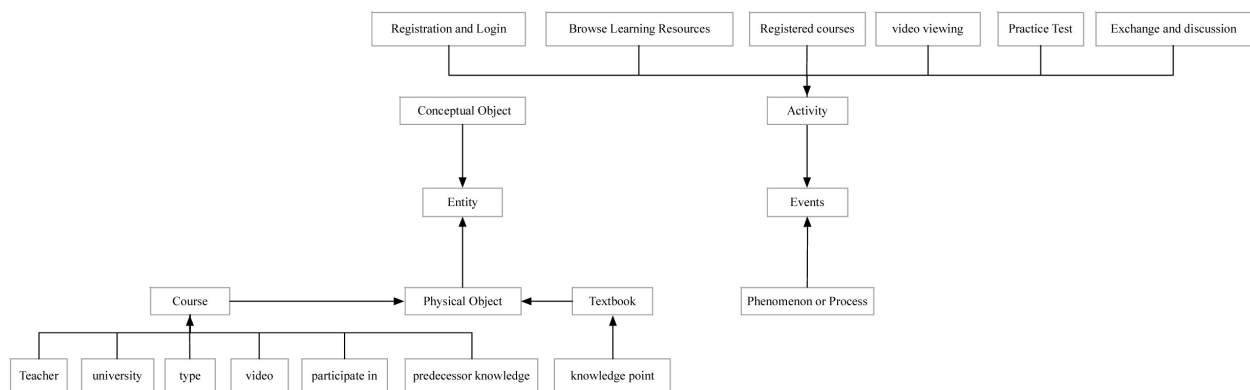The semantic type design of knowledge graph is shown in **Figure 1**.

Figure 1. The semantic type design of knowledge graph.

2.2. Semantic Relationship Design of Knowledge Graph

Semantic relations in the semantic network describe the relationships among all semantic types in the semantic network, that is, the links between semantic types, corresponding to the connections between nodes in the semantic network, that is, they can be regarded as the relationships among entities in the knowledge graph. Following the principle of high-level relationship reusability, semantic relations are composed of two categories: "Associated_with" and "Is_a". The semantic relations of "Associated_with" in the knowledge graph constructed in this paper consist of five categories: "Conceptually Related", "Physically Related", "Functionally Related", "Spatially Related" and "Temporally Related". The semantic relations of "Is_a" are the basis of dividing the conceptual hierarchy of knowledge graph, and each concept must have one or more superior concepts.

3. Database Design

3.1. Data Acquisition and Data Storage

The data is obtained from the MySQL database on edX platform, and store the required data as. csv files. After the knowledge points of each course are extracted, a. csv file is compiled according to the experts' opinions, and the data is sorted according to the construction requirements of knowledge graph.

The sorted file is imported into the Neo4j to complete the construction and storage of the entities and relationships of the knowledge graph.

3.2. Neo4j Design

As a knowledge storage database, Neo4j keeps a complete knowledge graph. There are seven types of nodes in the designed knowledge graph model. The design results are shown in **Table 1**.

For a designed knowledge graph model, there are nine types of relationship designs between nodes.

1) "Contain": This relationship connects knowledge point entities with knowledge point entities. Refers to which knowledge points are included in this knowledge point.

2) "Next": This relationship connects knowledge point entities with knowledge point entities. It indicates that the next knowledge point pointed by this knowledge point can be used to obtain the subsequent knowledge point set of this knowledge point. In the same way, with the help of this relationship, the precursor knowledge point set of this knowledge can be obtained by obtaining which knowledge points the knowledge point is pointed to. According to the dependency between knowledge points, the learning path planning within the course can be carried out. Moreover, when learners encounter bottlenecks in their learning, it is easy to get the precursor knowledge points of the knowledge points that learners have not mastered well, so as to update the planned learning path.

Table 1. Node design of Neo4j.

Node	remarks
knowledge	Knowledge point entity, which stores the knowledge points possessed by professional courses. Attribute value: "id" is the unique identifier of knowledge point entity, and "name" indicates the name of knowledge point.
course	Course entity, which store courses on MOOC platform. Attribute value: "id" is the unique identification of the course entity, and "name" indicates the course name.
teacher	Teacher entity, which stores the lecturers on MOOC platform. Attribute value: "id" is the unique identifier of the teacher entity, and "name" indicates the teacher's name.
university	University entity, which stores the universities on the MOOC platform. Attribute value: "id" is the unique identification of the university entity, and "name" indicates the name of the university.
type	Type, which stores the course categories existing on MOOC platform. Attribute value: "id" is the unique identifier of the category entity, and "name" indicates the category name.
video	Video entity, which stores videos of courses on MOOC platform. Attribute value: "id" is the unique identifier of video entity, "video" indicates the video name, and "link" indicates the link address of the vide
question	Test questions entity, which stores the exercises of courses on MOOC platform. Attribute value: "id" is the unique identification of the test questions entity, "name" indicates the name of the test question, and "link": indicates the link address of the exercise.

3) "Next_course": This relationship links the course entities with the course entities. The relationship is obtained from the prerequisite courses in the course information on MOOC platform. When planning the learning path between courses, it is used to judge whether the learner has finished learning the courses in the prerequisite course set of recommended courses. If there are courses in the prerequisite course set of recommended courses that the learner has not taken, the courses in the prerequisite course set are first planned into the learning path and recommended to the learner. Make the learning path recommended to learners more professional.

4) "Knowledge_video": This relationship connects knowledge entities with video entities. It can be used to obtain the video set corresponding to a certain knowledge point and the knowledge point set corresponding to a certain video, which is convenient for learners to know which videos correspond to poorly mastered knowledge points and which videos correspond to poorly mastered knowledge points, which is convenient for planning and updating the learning path in the course.

5) "Question_knowledge": This relation connects the question entities and the

knowledge entities. It can be used to judge the knowledge points information corresponding to the test questions. When learners make mistakes in exercises after class, it is convenient for learners to judge which knowledge points they have not mastered well through the wrong questions, and it is convenient for updating the learning path.

6) "Contain_video": This relationship connects the course entities with the video entities. Refers to the video set corresponding to the course and which course the video belongs to.

7) "Teach": This relationship connects the curriculum entities with the teacher entities. It refers to which teacher taught the course.

8) "University": This relationship connects the curriculum entities with the university entities. Refers to which university the course is offered.

9) "Type": This relationship connects the course entities and the category entities. Refers to which category the course belongs to.

4. Learning Path Planning

4.1. Learning Path Planning between Courses

After the knowledge graph is constructed, according to the historical data of the courses that learners have studied on edX platform, the learning path planning between courses is completed.

In the knowledge graph, if there is a "next_course" relationship between course entity A and course entity B, it is represented as "course A" next_course "course B", then course A is called the pre-course of course B, and it can also be proved that course B is the post-course of course A. The set of all the pre-courses of a course is called the pre-course set of the course, that is, the set of all courses whose "next_course" relationship points to course B, which is called the pre-course set of course B. The set of all the post-courses of a course is called the post-course set of the course, that is, the set of all the courses pointed by the course A with the relationship of "next_course" is called the post-course set of the course A. According to the order of course learning, if learners want to learn from course B, they need to make sure that the pre-course set of course B is empty, or that the pre-course set of course B have been studied.

1) Basic course: In the knowledge graph, if the pre-course set of a course is empty, this course is called a basic course.

2) Completed Course Set: A set of all courses that the learner has studied on MOOC platform, which is called the completed course set of that learner.

3) Uncompleted course set: the set of all courses that a learner has not learned on MOOC platform, which is called the learner's uncompleted course set. Divide the courses existing in the Uncompleted course set into the inactive course set and the activated course set. If the course is a basic course or the learner has completed all the courses in the pre-course set of the course, then the set composed of such courses is called the activated course set. In the uncompleted course set, the set of courses that do not belong to the activated course set is

called the inactive course set.

On the edX platform, taking the learner's completed course set as the seed set in RippleNet algorithm, combining with knowledge graph to recommend courses, and based on the learner's activated course set, the learner's learning path planning between courses is completed.

At the same time, the introduction of RippleNet algorithm makes the recommended results in the learning path planning between courses not only diverse, but interpretable (Wang, Zhang, Wang, Miao, & Guo, 2018). Traditional collaborative filtering recommendation algorithms usually interpret recommendation results based on community tags, social networks, or phrase emotions. RippleNet algorithm is based on the entities and relationships in the knowledge graph to recommend for learners.

4.2. Learning Path Planning within the Course

According to the knowledge graph of the knowledge points involved in the course, the learning path planning of learners in the course is given. According to the learners' answers to after-class exercises, the learning path in the course is updated.

1) Predecessor and successor knowledge points: in the knowledge graph, if there is a "next" relationship between knowledge point entity A and knowledge point entity B, it is represented as "knowledge point A" <u>next</u> "knowledge point B" means that knowledge point A is the predecessor of knowledge point B, which also means that knowledge point B is the successor of knowledge point A. In the same way as the pre-course set, the set of knowledge point entities, which the "next" relationship points to knowledge B, is called the predecessor knowledge point set of knowledge B. The set of knowledge point entities pointed by knowledge entity A with the relationship of "next" is called the successor knowledge point set of knowledge entity A.

2) Initial knowledge point: If the predecessor knowledge point set of knowledge point C is empty and has a "contain" relationship with other knowledge points, knowledge point C is the initial knowledge point.

3) Terminal knowledge point: If the successor knowledge point set of knowledge point D is empty and has a "contain" relationship with other knowledge points, knowledge point D is the terminal knowledge point.

There is a certain relationship between knowledge points, which can be divided into three categories:

1) Father-son relationship: it reflects the whole-part relationship between knowledge points. If there is a "contain" relationship between knowledge point A and more than one knowledge point entity, knowledge point A is the parent knowledge point of all knowledge point entities it contains, and each knowledge point entity contained in knowledge point A is called the child knowledge point of knowledge point A.

2) Precedence relationship: The precedence relationship between knowledge entities is derived from the concepts of the predecessor knowledge points and

the successor knowledge points of knowledge entities. The order of knowledge points is used to judge the depth of learners' mastery of the course. According to the order of knowledge point learning, if knowledge point A is in the predecessor knowledge point set of knowledge point B, learners can learn the successor knowledge point B only on the basis of learning knowledge point A. Therefore, if you want to learn knowledge point B, you must first learn the predecessor knowledge point set of knowledge point B.

3) Parallel relationship: refers to two or more knowledge points at the same level. There is no dependency relationship between them, and there is no whole-part relationship. The parallel relationship between knowledge points is used to judge the breadth of learners' mastery of the course.

On the edX platform, according to the learners' answers to the exercises, judge their mastery of the course, and give feedback to the learners' mastery of the course knowledge points in a visual way. Learners can check the knowledge points in the predecessor knowledge points set and the successor knowledge points set by inputting knowledge points on the platform. According to the answers of the exercises corresponding to the knowledge points, we can judge the reasons why learners don't know the knowledge points and which knowledge points to learn next, and update learning path planning in the course.

4.3. Application of Visualization Technology

At present, after completing the course study on the MOOC platform, learners do not get the feedback information of the learning effect, such as which knowledge points they have studied are not firmly grasped, and which knowledge points should be reserved before learning this knowledge point. Learners do not know how much they have mastered, and updating the learning path without distinguishing the differences of learners' learning situation will make learners learn a lot of knowledge that is inconsistent with their learning ability and needs. Visualization technology allows learners to know more about their mastery of what they have learned, and at the same time, understand the reasons for the updated learning path. Visualization and knowledge graph technology add interpretability to the planning of learning path (Bauman, Bing, & Tuzhilin, 2017).

The visual interaction content on MOOC platform is as follows. Aiming at the exercises that learners answer incorrectly, feedback the knowledge points corresponding to the wrong questions for learners in a visual way. Learners can independently input the knowledge points they want to query, and the platform feeds back the predecessor knowledge points of the knowledge points and the videos corresponding to the successor knowledge points to the learners. Feedback the updated learning path of course to learners, and learners can make their own choices.

5. Application of Knowledge Graph

On MOOC platform, the application of knowledge graph includes is as follows.

Make use of the internal relationship between knowledge points in knowledge graph to plan the learning path for learners in the course. Make use of the internal relationship between courses in knowledge graph and RippleNet algorithm to plan the learning path between courses for learners. Correspond the knowledge points in the knowledge graph to the exercises on the edX platform, and combine the Echarts visualization tool, and visually feedback the knowledge points corresponding to the wrong questions to the learners. Based on PyCharm development environment, Py2neo development framework and Echarts visualization tool, learners can interact with knowledge graph and query knowledge information. It provides a reasonable explanation for the updated learning path.

6. Conclusion

Based on the edX platform, this paper designs the knowledge graph model by combining expert method and reference method, select Neo4j as the knowledge storage database of knowledge graph; selects Echarts as the visualization tool of Neo4j, finally realizes the visual search and display of knowledge points on MOOC platform, and gives learning feedback and update according to the learning situation of learners. When planning the learning path through the designed knowledge graph, Ripplenet algorithm is introduced to enrich the diversity of courses in the learning path planning between courses. Professional learning path planning is provided for learners to improve the learning efficiency of online learning.

This paper lies in the introduction of RippleNet algorithm while learning path planning based on knowledge graph, which enriches the diversity of curriculum recommendation in learning path planning. Through data visualization technology, learners are given learning feedback on MOOC platform, so that learners can understand the knowledge points and the reasons for learning path update. The introduction of knowledge graph technology and data visualization technology is conducive to improving the accuracy, professionalism and interpretability of learning path planning; It is helpful to reasonably enhance learners' interest and enrich the diversity of recommendations.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Bauman, K., Bing, L., & Tuzhilin, A. (2017). Aspect Based Recommendations: Recommending Items with the Most Valuable Aspects Based on User Reviews. *The 23rd ACM SIGKDD International Conference.* https://doi.org/10.1145/3097983.3098170

Cao, Y. C., & Zuo, Y. L. (2018). An Analysis of the Current Situation of Web-Based Courses in China. *Technology and Innovation, No. 17,* 2.

Li, Y., Shao, Z., Wang, X., Zhao, X., & Guo, Y. (2018). A Concept Map-Based Learning Paths Automatic Generation Algorithm for Adaptive Learning Systems. *IEEE Access.*

https://doi.org/10.1109/ACCESS.2018.2885339

NCBI (2022). UMLS Reference Manual. *APA Divisions*.
http://www.ncbi.nlm.nih.gov/books/NBK9676/

Wang, H., Zhang, F., Wang, J., Miao, Z., & Guo, M. (2018). Ripplenet: Propagating User Preferences on the Knowledge Graph for Recommender Systems. *ACM.* *https://doi.org/10.1145/3269206.3271739*

Yang, X. H., & Zhuo, X. Z. (2017). Research on the Current Situation and Development Trend of Online Education in China: Based on the Analysis of 16 Online Education Platforms Such as Netease Open Class. *E-Education Research, 38,* 8.

Zhang, M. Y. (2019). Analysis and Research on Learning Paths in the Environment of Online Open Courses: Taking the Wisdom of the Analects of Confucius and Other Two Courses as Examples. Doctoral Dissertation. Shanghai: Shanghai Normal University.