

Gradual Study Advising with Course Knowledge Graphs

Junnan Dong[®], Wentao Li[®], Yaowei Wang[®], Qing Li[®], George Baciu[®], Jiannong Cao[®], Xiao Huang^(⊠)[®], Richard Chen Li[®], and Peter H. F. Ng[®]

Abstract. Knowledge graphs (KGs) have been actively studied for pedagogical purposes. To depict the rich but latent relations among different concepts in the course textbook, increasing efforts have been proposed to construct course KGs for university students. However, the application of course KGs for real study scenarios and career development remains unexplored and nontrivial. First, it is hard to enable personalized viewing and advising. Within the intricate university curricula, instructors aim to assist students in developing a personalized course selection pathway, which cannot be fulfilled by isolated course KGs. Second, locating concepts that are important to individuals poses challenges to students. Real-world course KGs may contain hundreds of concepts connected by hierarchical relations, e.g., *contain_subtopic*, making it challenging to capture the key points. To tackle these challenges, in this paper, we present GSA, a novel gradual study advising system based on course knowledge graphs, to facilitate both intra-course study and inter-course development for students significantly. Specifically, (i) we establish an interactive web system for both instructors to construct and manipulate course KGs, and students to view and interact. (ii) Conceptlevel advising is designed to visualize the centrality of a course KG based on various metrics. We also propose a tailored algorithm to suggest the learning path based on what concepts students have learned. (iii) Courselevel advising is instantiated with a course network. This indicates the prerequisite relation among different levels of courses, corresponding to the annually increasing curricular design and forming different major streams. Extensive illustrations show the effectiveness of our system.

Keywords: Study Advising \cdot Knowledge Graphs \cdot Graph Visualization

1 Introduction

Knowledge graphs (KGs) have emerged as a potent tool for enhancing pedagogical achievements [9]. They could effectively represent the unstructured knowledge from textbooks as triples [6], i.e., (*Relational query languages*, contain_subtopic, Procedural language), unveiling the latent connections between complex concepts and facilitating students' understanding rather than directly learning with the whole textbook. This particularly holds true and inspiring within the realm of college education. Recently, the course KG construction has gained much research attention [7]. Early studies propose deep learning methods to automatically construct KGs for education for their potential to illuminate complex relations among course concepts [2]. Another group of methods leverages online platforms [3,13], e.g., Wikipedia and MOOC, to enhance the concept extraction performance by linking the entities in the textbook with additional knowledge sources. However, though substantial efforts have been invested in constructing course KGs to capture the nature of concepts in course textbooks, they primarily focus on the construction and visualization of course KGs but fail to adequately address the practical difficulties encountered during students' utilization. A critical gap persists in bridging the course KGs with real-world study scenarios and subsequent career development. As educators seek innovative ways to empower university students with not only a good command of the course itself but also forming their own major stream through course selection, existing methods are not applicable to the downstream situation, also the related research remains limited and unexplored.



Fig. 1. A real-world example of the course knowledge graph constructed for 'Database Systems'. It contains 96 triples connected by *contain_subtopic*.

Motivated by this, we investigate a tailored study advising system based on course knowledge graphs for both *intra-course* learning and *inter-course* career trajectories. However, this task is challenging for two major challenges.

First, isolated course KGs can hardly enable personalized viewing and advising. Within the intricate university curricula, instructors aim to assist students in developing individualized course selection pathways. To be associated with the aim of *inter-course* advising, isolated course KGs should be connected with each other in a logical way. An intuitive solution is shown in Fig. 2, which directly



Fig. 2. An intuitive solution for students to obtain the overall academic picture by directly combining isolated course KGs.

combines all the course KGs in one, it is even harder to be interpreted. This makes study advising with course KGs for personalized course selection, as well as broader career planning, a nontrivial task. *Second*, it is inefficient for students to locate concepts that are important. In real-world scenarios, course KGs may contain hundreds of concepts connected by hierarchical relations, e.g., *contain_subtopic*, identifying crucial concepts within the complicated graph makes course KG-based study advising non-trivial. In Fig. 1, we visualize a course KG constructed for 'Database Systems', which is one of the smallest course KGs existing with merely 96 triples, in the form of a mind map for a clear illustration 1a. Despite the conciseness of this mind map, it is obviously time-consuming to read all the concepts for *intra-course* learning. The situations are significantly more complicated within the commonly used knowledge graph visualization with Neo4j database in Fig. 1b.

To this end, we present a novel approach, i.e., Gradual Study Advising (GSA), which leverages course KGs to effectively facilitate both intra-course comprehension and inter-course career development for university students. Specifically, (i)we first establish a basic interactive web-based system for university instructors to create and manipulate the suitable course KGs, granting them the tools to construct dynamic and informative graphs; (ii), we design concept-level advising, a novel visualization mechanism, corresponding to the *intra-course learning*, that quantifies the centrality of concepts within course KGs using a spectrum of metrics. Moreover, a substantial algorithm is proposed to tailor learning pathways based on students' acquired knowledge; (iii) A course-level advising is empowered by constructing a course network that uncovers the prerequisite relations among courses, enabling students to chart distinct career streams. This is inspired by the computer science curricula development in the Department of Computing, The Hong Kong Polytechnic University. For one CS freshman, four years are systematically designed from fundamental (year 1), broadening (year 2), and strengthening (year 3) to Specialization (year 4).

In general, our contributions are summarized below.

- We propose a new paradigm for gradual study advising with course knowledge graphs for intra-course learning and inter-course development.
- An interactive system is developed to facilitate both instructors to convey course knowledge and students to utilize it.
- Sufficient illustrations are provided to illustrate the effectiveness of our proposed system.

2 Functional Foundations for Interactive Web System

In this section, we introduce how we preliminarily prepare an interactive system for instructors to manipulate and publish a course KG to students, as well as for students to seek study advising. We first elaborate on the initialization of course KG construction based on textbooks in Sect. 2.1 and the tailored online manipulator for instructors in Sect. 2.2.

2.1 Course Knowledge Graph Construction

We undertake a comprehensive revision of the ontology of course KGs, which has been specifically tailored to better accommodate educational purposes in [7]. Our primary focus is the construction of 'contain_subtopic', which provides a clear picture of conceptual relationships hierarchically, thereby facilitating students' intr-course understanding and the establishment of links between different concepts. For each course, we began with a list of'seed entities', which are the core concepts identified by experienced educators that underpin the respective curriculum. Primary sources such as textbooks and Wikipedia were harnessed to build the course KGs, with a distinct emphasis on enlarging the graph centered around these seed entities. Then, we use the relation extraction model to output new triples, the input is a section of the textbook relevant to the seed entities parsed from a PDF file. For the extracted triples, we employed the seed entities for further filtering as we score the importance of each triple, ensuring that each EKG encapsulates the most essential information, thereby minimizing redundancy.

In particular, for computer science education in year 1: we identified approximately 20 distinct courses for our course KGs, each containing 10 to 20 seed entities. After the extraction, each course includes between 50 to 100 triples and an equivalent number of entities. Furthermore, we proceeded to merge and organize the different subgraphs obtained for each subject, which mainly involved entity alignment and redundancy checks. This step was primarily accomplished through natural language processing algorithms and manual rules.

2.2 Course Knowledge Graph Management for Instructors

As a content management module for course KGs, in this section, we aim at providing management tools that support instructors' common activities in maintaining a course KG. Several managerial services are enabled by the included visualization interface and version control system. The visualization interface provides an orientation to the content and the relationship between them to users. The version control system expedites multiple managerial tasks that concern access right control, collaboration, and communication.



(a) A workspace for course KG manipulation.

(b) Managerial tools for course materials and links.



(c) Highlights of unstaged modification in the workspace. Unstaged triple creations are represented by solid lines. Unstaged triple removals are represented by solid broken lines. All staged triples are suppressed and turned semi-transparent.

Fig. 3. A simple illustration of the functionalities of the course KG manipulator.

Course Knowledge Graph Manipulator. Given the rapid changes in teaching targets based on students' feedback, to fulfill a convenient manipulation of course KGs, an integrated workspace is developed for instructors to easily manipulate course KGs and course material. Following the prevailing work [1], GSA provides a workspace interface (see Fig. 3), where instructors can manipulate the knowledge graph with the elementary operation, e.g., add or delete. For creation, the relation in a triple is defaulted as 'contain_subtopic'. Advanced operations, e.g. removal of nodes that are unreachable from the course node, are also available. In our manipulator, concepts are associated with the triples, each time an addition should take effect as a head/tail node of a triple, depicted in Fig. 3. Specifically, for new concept addition, instructors first input the name of a new concept, i.e., 'Newly Added Concept' and click the corresponding tail entity and click 'LINK' to generate triples in the graph. While for a new connection

between existing concepts, manipulations can be easily done by either clicking the existing concept or choosing from the list.

While for managing the auxiliary learning materials, course material of an arbitrary concept node could be added through a pop-up window (see Fig. 3b). Information, such as providers, textual description, and the URL of the material, e.g., from Wikipedia, is stored by **GSA**, which also facilitates intra-concept learning.

Additionally, the workspace interface takes version control into consideration. The unstaged modification would not be tracked by the version control system. They are considered undocumented and volatile. Unstaged modification could be highlighted in an edit mode (see Fig. 3c). A workspace with unstaged modification could be set visible to students. This might be convenient in situations where the instructor has to expeditiously publish changes of the syllabus without the time to make precise remarks and tags for version control purposes. The downside of such an experience is that it may discourage instructors to proceed with a formal versioning process. As a countermeasure, highlighting unstaged remarks facilitates users when they want to track unstaged changes and reminds users to perform proper versioning and documentation tasks.

Algorithm 1. Algorithm of cross-tab synchronization using local Storage
1: Tab 1 for Instructor A:
2: Set the data to be shared using local storage.
3: End Tab 1
4:
5: Tab 2 for Instructor B:
6: while listening for changes in local storage do
7: If a change event occurs:
8: if the changed key is 'sharedData' then
9: Retrieve the new value from local storage.
10: Display the updated value with Tab2: [received data].
11: end if
12: end while
13: End Tab 2

Cross-Tab Synchronization. In consideration of the collaborative manipulation scenario among different instructors that may be responsible for the same course, they may work together in the workspace to create, edit, and manage course content. With cross-tab synchronization, changes made by one instructor are immediately reflected in all open tabs or instances where the course knowledge graph is being viewed or edited. This real-time collaboration ensures that instructors can see each other's changes without delays, fostering efficient teamwork.

We explain this process in Algorithm 1. Specifically, in the instructor A's tab, while making changes to the course KG, the instructor updates the graph

according to their actions. After each update, a synchronization event is triggered to notify other tabs that changes have been made. In other tabs where the same course knowledge graph is being viewed or edited, a continuous listening loop monitors synchronization events. Upon detecting a synchronization event, these tabs receive the updated course knowledge graph data and update their displayed knowledge graphs to mirror the changes made by the instructor.

This process ensures that all tabs displaying the course KG remain in sync, providing instructors with a cohesive and real-time collaboration environment. The cross-tab synchronization approach for instructors' manipulation of course KGs offers several benefits:

- Real-time Collaboration: Instructors can collaboratively work on the same course content, seeing each other's changes in real-time.
- Seamless Experience: Changes made by one instructor are immediately reflected across all tabs, eliminating confusion or discrepancies.
- Enhanced Productivity: Instructors can focus on content creation and manipulation without interruption or manual updates.



Fig. 4. A dual framework for study advising from course level and concept level.

3 Approach: GSA

As illustrated in Fig. 4, we propose a gradual study advising framework that consistently integrates both intra-course learning and inter-course advising. We aim to first guide the students to form their personalized major stream, as well as the career development through inter-course advising, and then dig deeply into particular courses with the intra-course learning module.

3.1 Inter-course Advising

Course Network Construction. In order to provide gradual inter-course advising, we have developed a sophisticated course network that interconnects all the courses through the "prerequisites of" relationship. Instead of combining all the course KGs together, m this innovative approach significantly reduces the graph size and highlights the progressive course-level relations.

We draw inspiration from the curricular design framework utilized by the Department of Computing at The Hong Kong Polytechnic University. Following a meticulously structured pathway, CS freshmen undergo a four-year journey, progressing systematically from fundamental (year 1) to broadening (year 2), strengthening (year 3), and culminating in specialization (year 4).



Fig. 5. The intra-course advising provides visualization of prerequisite relations among courses. Students can toggle the visualization. GSA also allows them to open a list that includes all textual paths that contain particular courses.

Relational Path Finder. Following this transformative approach, students are empowered to navigate the course network according to their year level, aligning their course selections with their career aspirations and major streams. For instance, in the Department of Computing at the Hong Kong Polytechnic University, those who aspire to become fin-tech experts can strategically choose courses in a coherent sequence. In year 1, they could select fundamental knowledge, easing their transition to university studies with an introduction to Scheme. In the second year, they could select the courses, of which they have taken the prerequisites, to acquire broad computing skills, along with rudimentary concepts of economics, accounting, and finance. The third year is dedicated to continuing to strengthen core competencies, encompassing software engineering, systems security, and a selection of computing or finance electives. Finally, in the fourth and final year, students specialize in areas like artificial intelligence, machine learning, and pattern recognition, as well as emerging fields like crowdfunding, e-finance, and e-payment systems. This curriculum-guided inter-course advising ensures a logical progression, enabling students to make informed and strategic choices in line with their evolving career goals.

To achieve this goal, we design a relational path finder where prerequisite relations between courses could be visualized through GSA. Visualization of these relationships is available in the homepage. Particularly, the visualization of the prerequisite relationship is illustrated in Fig. 5. Such visualization orients users

about the overall structure of the university program. For example, the prerequisite relationship among courses in this case could help students to decide what course they should select or review for future stream development.

3.2 Intra-course Advising

To tackle the challenges that course KGs are difficult to be utilized by students given the number of concepts and complicated connections, in this subsection, we employ different centrality metrics such as degree centrality, and PageRank, to clearly depict the centrality of one course KG, providing students with a dynamic visual representation toolbar that could highlight the significance of concepts within the course KG. These visualizations enable students to quickly grasp the core ideas and critical nodes within the knowledge graph, promoting efficient learning. The centrality metrics illuminate nodes with high connectivity, bridge nodes that connect disparate areas, and influential nodes that carry substantial importance. Through interactive and intuitive visualizations, students can identify pivotal concepts, explore relationships, and navigate the course KG's complexity with ease. These methodologies not only facilitate the rapid acquisition of key knowledge but also empower students to comprehend the interconnections that underlie the course content, fostering a deeper understanding of the subject matter.

In this paper, we showcase two aspects of centrality visualization by employing 'Degree' and 'PageRank'.



Fig. 6. The centrality visualization based on degrees.

Degree Centrality. In the realm of intra-course learning, the visualization of centrality through the lens of degree centrality emerges as a powerful tool. The concept of degree centrality brings forth a structured approach to understanding the pivotal nodes within a course's knowledge graph. This visualization technique is underpinned by the calculation of the degree of a concept, which reflects

its connectedness to other concepts within the graph. In essence, the degree centrality $deg(e_i)$ of a concept e_i , encapsulates the sheer number of relationships linked to it. Mathematically articulated as:

$$deg(e_i) = Len_{r \in N_{e_i}}(r), \tag{1}$$

where N_{e_i} is the one-hop neighbor triples centered by concept e_i , and r represents the relation appears in N_{e_i} . This metric holds profound significance. By quantifying the number of edges or relationships incident to an entity, degree centrality provides a quantitative representation of its influence and importance within the graph. In Fig. 6, we visualize the central concepts with colors from shallow to deep.

In the context of intra-learning with course KGs, this visualization approach serves as a compass, guiding learners toward the concepts that play a pivotal role in shaping their understanding of key concepts. Through degree centrality visualization, students gain an intuitive grasp of the central components that underpin the course's knowledge structure, enhancing their ability to navigate and comprehend complex subject matter.

PageRank Centrality. Similar to degree centrality, we visualize the result of PageRank in Fig. 7.



Fig. 7. An alternative visualization of centrality based on the 'PageRank' metric. The adapted PageRank equation for a knowledge graph, considering entities and their relationships, is formulated as follows:

$$PR(E_i) = (1 - d) + d \times \sum_{e_j \in N_{e_i}} \frac{PR(e_j)}{|L(e_j)|}$$
(2)

where $PR(E_i)$ is the PageRank score of concept e_i . $e_j \in N_{e_i}$ is the set of neighbor concepts that link to entity e_j . d is the damping factor, a value between 0 and 1, representing the probability that a student follows a connection rather than jumping to another random concept. Comparing 'Degree' and 'PageRank' as centrality metrics, we adopt distinct approaches to visualize the importance of nodes from different views. Degree centrality focuses on the straightforward notion of connectedness. By counting the number of relationships linked to an entity, 'Degree' identifies nodes with high interaction and participation within the graph. Intuitively it is computationally simple for rendering and quickly identifying heavily connected concepts. However, it might overlook concepts that are indirectly influential due to their position in the graph. While 'PageRank' introduces a more nuanced perspective by considering not only the number of relations but also the quality of those connections. This metric reflects the importance of one concept which is not solely determined by its own degree but also by the importance of concepts linking to it. This algorithmic approach accounts for the graph's structure and provides a more sophisticated understanding of influence.

In general, for intra-course learning in our GSA, different choices of centrality depend on the learning objectives and the nature of the course KGs. We provide a balanced approach that might involve employing various metrics, leveraging all their insights for a more refined understanding of centrality and influence.

3.3 Concept Learning Path Recommendation

In addition to the centrality visualization, for advanced intr-course learning with course KGs, we also propose a tailored recommendation algorithm that still remains unimplemented. As shown in Fig. 4, given a course KG, denoted as C, we would like to do personalized and time-sensitive recommendations based on the semester teaching schedule and how well the student grasps the current progress. We design an expectation score that evaluates the importance of one concept e_i that should be recommended for the student s to preview or review based on the schedule. The proposed recommendation is formulated as a ranking problem. **GSA** will first calculate the expectation score $E(e_i, s)$ with a multiplication of: (i) the importance score of concept e_i , $deg(e_i)$; (ii) the relatedness of the concept R_{N_i} considers how many of the subtopics/prerequisite concepts have been taught according to current progress, the equations are derived as:

$$E(e_i, s) = deg(e_i) \times R_{N_i} \tag{3}$$

Within each interaction, **GSA** will first traverse the academic calendar and the semester schedule for particular courses. Then a personalized study progress will be retrieved based on the timestamp which indicates the courses and the concepts in each course that the student has taken. Finally, we calculate the expectation score for each candidate concept and orderly sort them by rankings. Those with higher scores are expected to be recommended to students for their review/preview subject to the temporal state.

4 Related Work

4.1 Educational Knowledge Graph Construction

Knowledge graphs have gained traction in education for enhancing the representation and navigation of educational content. In early studies, KnowEdu [2] proposes to combine deep learning and rule mining methods, i.e., GRU and p-Apriori, to extract knowledge from internal system data and evaluate it with two human experts' annotation on all entities and relations. Recently, efforts have been made to leverage various e-learning platforms, e.g., Wikipedia and MOOC. EduKG [13] constructs knowledge graphs from educational resources to aid content recommendation. MOOCKG [3] links course concepts with external knowledge sources to enrich course content. In this paper, we bridge the gap between educational KG construction and effective academic advising for students with our GSA system.

4.2 Study Advising

The field of study advising has witnessed substantial advancements in recent years [10]. Several notable studies and systems have contributed to the understanding and implementation of study advising [1,4,5]. Early advisors have typically relied on face-to-face interactions to provide guidance to students. These interactions often involve discussions about course selection, career paths, and academic progress. Garton et al. [11] emphasize the importance of personalized interactions in their study on student perceptions of academic advising. Sweker et al. [12] investigate the importance of the number of meetings between advisors and first-generation students. Recently, with the advancement of technology, various digital tools have been developed to enhance study advising processes. Online platforms, such as advising portals and educational planning software, have been designed to facilitate communication between students and advisors. MacDonald et al. [8] suggest the power of distant advising through online platforms, revealing the impact of technology-enhanced advising platforms on student engagement and satisfaction. In our GSA, we empower instructors with an intelligent interactive system to manipulate the course knowledge graph and automatically provide both intra-course and inter-course advising.

5 Conclusions and Future Work

In this paper, we present a novel Gradual Study Advising system, i.e., GSA that emerges as a pioneering solution to bridge the gap between course knowledge graph (course KG) construction and students' practical needs of learning and career development with course KGs. The fundamental role of course KGs in reshaping pedagogical approaches has led us to formulate a tailored advising system that integrates both intra-course learning and inter-course career. Specifically, we empower university instructors with an interactive web-based platform that allows them to craft dynamic and informative course KGs, ensuring the quality of comprehensive course KGs. Moreover, our innovative concept-level advising mechanism transforms course KGs into easily understandable visualizations, by quantifying concept centrality through a range of metrics. This enables students to comprehend the hierarchy of concepts within course KGs and paves the way for personalized learning pathways. Finally, we extend our approach to inter-course advising, establishing a course network that uncovers prerequisite relations among courses, and guiding students to chart distinct career paths inspired by the progressive curricula design in the Department of Computing at the Hong Kong Polytechnic University.

In future work, we will continue implementing the auto-advising on concept learning path recommendation based on personalized study achievements. By incorporating adaptive learning algorithms, GSA will dynamically tailor learning pathways and career trajectories based on individual student preferences and their taken concepts, as well as courses, according to the semester schedule. This would amplify the effectiveness of our system in guiding students toward their academic and professional goals.

Acknowledgement. This work was supported in full by the Hong Kong Polytechnic University, Project of Strategic Importance (project number: P0036846).

References

- Chan, J.C.K., et al.: Intelligent instructional design via interactive knowledge graph editing. In: González-González, C.S., et al. Learning Technologies and Systems. ICWL SETE 2022 2022. LNCS, vol. 13869, pp. 41–52. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-33023-0_4
- Chen, P., Lu, Y., Zheng, V.W., Chen, X., Yang, B.: KnowEdu: a system to construct knowledge graph for education. IEEE Access 6, 31553–31563 (2018)
- Dang, F., Tang, J., Li, S.: MOOC-KG: A MOOC knowledge graph for crossplatform online learning resources. In: 2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC), pp. 1–8. IEEE (2019)
- Hutson, B.L.: The impact of an appreciative advising-based university studies course on college student first-year experience. JARHE 2(1), 4–13 (2010)
- Junco, R., Mastrodicasa, J.M., Aguiar, A.V., Longnecker, E.M., Rokkum, J.N.: Impact of technology-mediated communication on student evaluations of advising. NACADA 36(2), 54–66 (2016)
- Kejriwal, M.: Domain-Specific Knowledge Graph Construction. SCS, Springer, Cham (2019). https://doi.org/10.1007/978-3-030-12375-8
- Li, W., et al.: Constructing low-redundant and high-accuracy knowledge graphs for education. In: González-González, C.S., et al. Learning Technologies and Systems. ICWL SETE 2022 2022. LNCS, vol. 13869, pp 148–160. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-33023-0_13
- 8. MacDonald, M., Garcia, B.W.: Advising the online student: a breakout of advising frequency, preferences, and satisfaction of online students. IJDLA **23**(3), 13 (2020)
- Rizun, M., et al.: Knowledge graph application in education: a literature review. Acta Universitatis Lodziensis. Folia Oeconomica 3(342), 7–19 (2019)

- 10. Steele, G.E.: Student success: academic advising, student learning data, and technology. NDHE **2018**(184), 59–68 (2018)
- 11. Suvedi, M., Ghimire, R., Millenbah, K., Shrestha, K.: Undergraduate students' perceptions of academic advising. NACTA J. **59**(3), 227–233 (2015)
- Swecker, H.K., Fifolt, M., Searby, L.: Academic advising and first-generation college students: a quantitative study on student retention. NACADA 33(1), 46–53 (2013)
- 13. Zhao, B., et al.: EDUKG: a heterogeneous sustainable k-12 educational knowledge graph. arXiv preprint arXiv:2210.12228 (2022)