Prerequisite Relation Learning for Concepts in MOOCs  
Liangming Pan, Chengjiang Li, Juanzi Li and Jie Tang  
Knowledge Engineering Group, Department of Computer Science and Technology, Tsinghua University

---

Introduction

- The prerequisite relation represents the learning dependency between knowledge concepts and is the cornerstone for designing curriculum in schools and universities.
- Our main contributions include:
  - The first attempt to detect prerequisite relations among concepts in MOOCs
  - Proposal of a set of novel features that utilize contextual, structural and semantic information in MOOCs to identify prerequisite relations
  - Design of three useful datasets based on real courses of Coursera to evaluate our method

Approach

- **Problem Definition**
  - **Input**: $D = \{c_1, \ldots, c_n\}$, where $c_i$ is one course.
  - $C = \{v_1, \ldots, v_m\}$, where $v_i$ is the $i$-th video of course $c$.
  - $U = \{k_1, \ldots, k_p\}$, where $k_i$ is the $i$-th sentence of the video texts.
  - $K = K_1 \cup \cdots \cup K_n$, where $K_i$ is the set of course concepts in $c_i$.

- **Output**:
  - **Prediction Function** $PF(a, b) \in \{0, 1\}$, $a, b \in K$, the function predicts whether concept $a$ is a prerequisite concept of $b$.

- **Video Reference Distance**: If in videos where concept A is frequently talked about, the teacher also needs to refer to concept B for a lot but not vice versa, then B would more likely be a prerequisite of A.

- **Complexity Level Distance**: If two related concepts have prerequisite relationship, they may have a difference in their generality level. It means that one concept is more general while another one is more specific.

Results

- **Datasets**
  - Collect course video captions from Coursera with three different domains.
  - Extract candidate concepts from and label the candidates as “course concept” or “not course concept”
  - Manually annotate the prerequisite relations among the labeled course concepts.

- **Performances with different classifiers**

- **Comparison with baselines**

Conclusion

We conducted a new investigation on automatically inferring prerequisite relations among concepts in MOOCs. We precisely define the problem and propose several useful features from different aspects, i.e., contextual, structural and semantic features.