EDM 2016

Course Content Analysis: An Initiative Step toward Learning Object Recommendation Systems for MOOC Learners

Yiling Dai, Yasuhito Asano, Masatoshi Yoshikawa
Graduate School of Informatics
Kyoto University

Background

Rising needs for diversity in one’s knowledge

Much easier to ? challenge new subjects

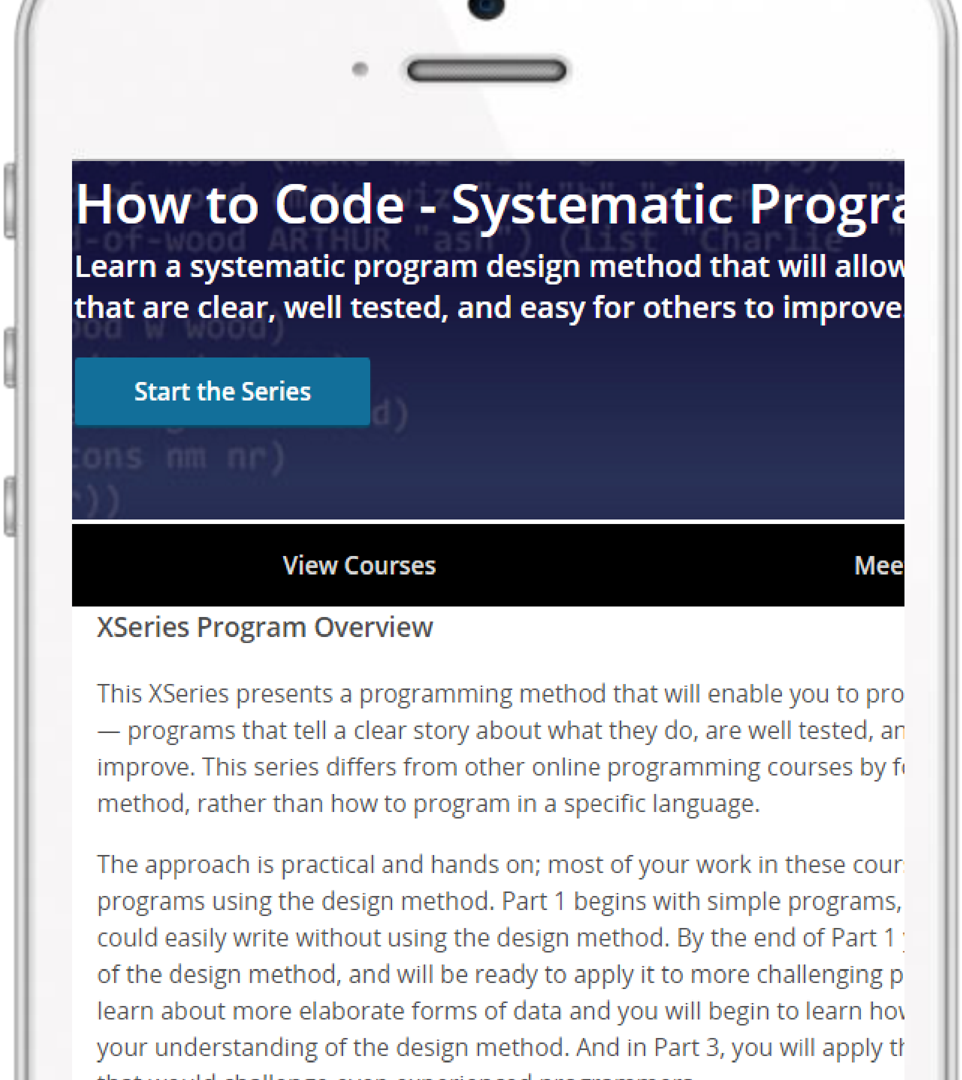
Background

For novices:

- No guidance
- Free to combine courses
- Similar courses and difficult syllabi

⇩

A supporting system may help!



How to Code - Systematic Progra

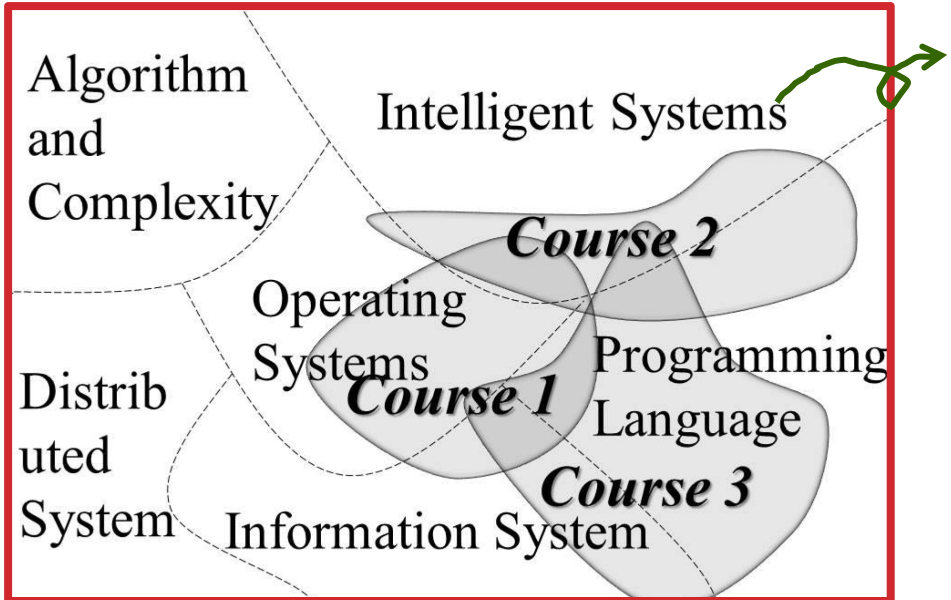
Learn a systematic program design method that will allow
that are clear, well tested, and easy for others to improve

XSeries Program Overview

This XSeries presents a programming method that will enable you to pro
— programs that tell a clear story about what they do, are well tested, an
improve. This series differs from other online programming courses by fo
method, rather than how to program in a specific language.

The approach is practical and hands on; most of your work in these cour
programs using the design method. Part 1 begins with simple programs,
could easily write without using the design method. By the end of Part 1 y
of the design method, and will be ready to apply it to more challenging p
learn about more elaborate forms of data and you will begin to learn how
your understanding of the design method. And in Part 3, you will apply th
that would challenge even experienced programmers.

Intuition



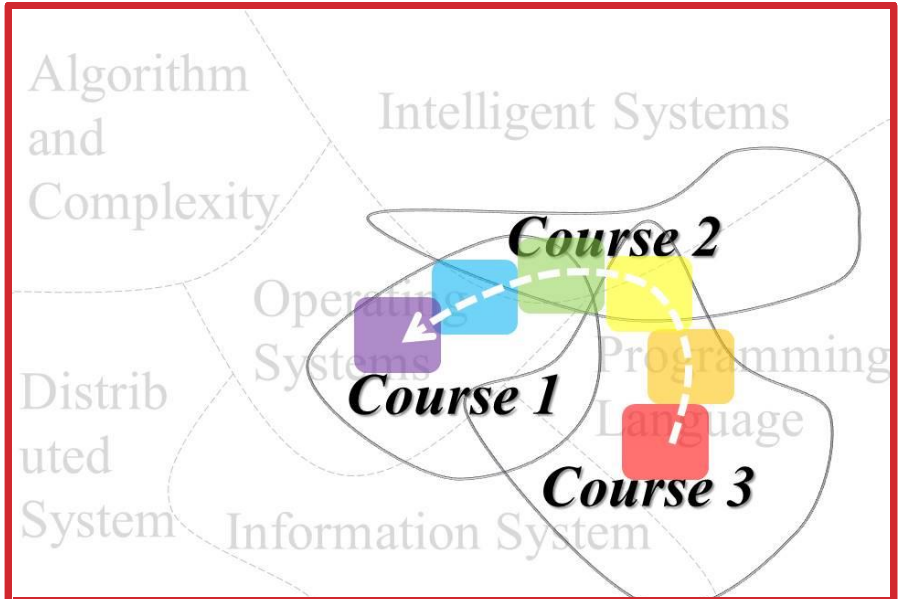
An image of the course map

Categories of domain knowledge

A course map

that presents the courses based on the knowledge they cover

Recommendation system

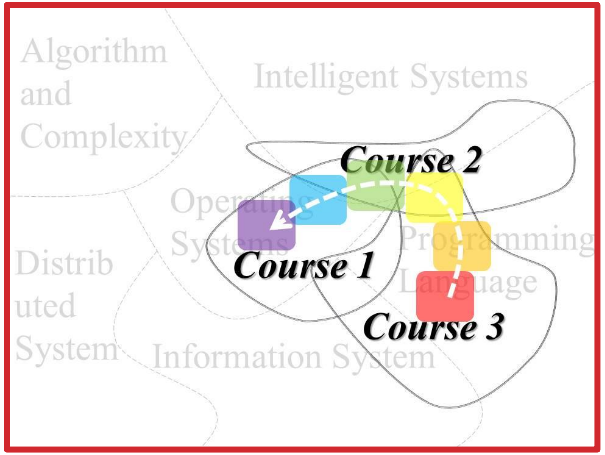


An image of the recommendation system

Advanced functions:

- Interact with the learner
- estimate his learning goals and ability
- Provide a path of learning objects (i.e. course lectures).

Task recognition



Final goal

Subtasks:

- **Course analysis**
  - Domain knowledge representation
  - Course content analysis
- **Learner analysis**
  - Background knowledge estimation
  - Learning goal recognition
  - …

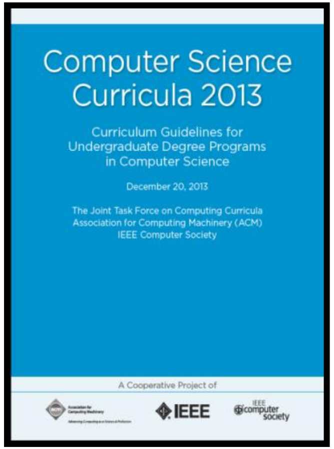
Course content analysis /Basic idea

Target:

Locating a course to the domain knowledge base

```
Course 1
--------------------------
--------------------------
--------------------------
--------------------------
--------------------------
--------------------------
--------------------------
-------
```

Course syllabus



Computer Science Curricula 2013

Curriculum Guidelines for Undergraduate Degree Programs in Computer Science

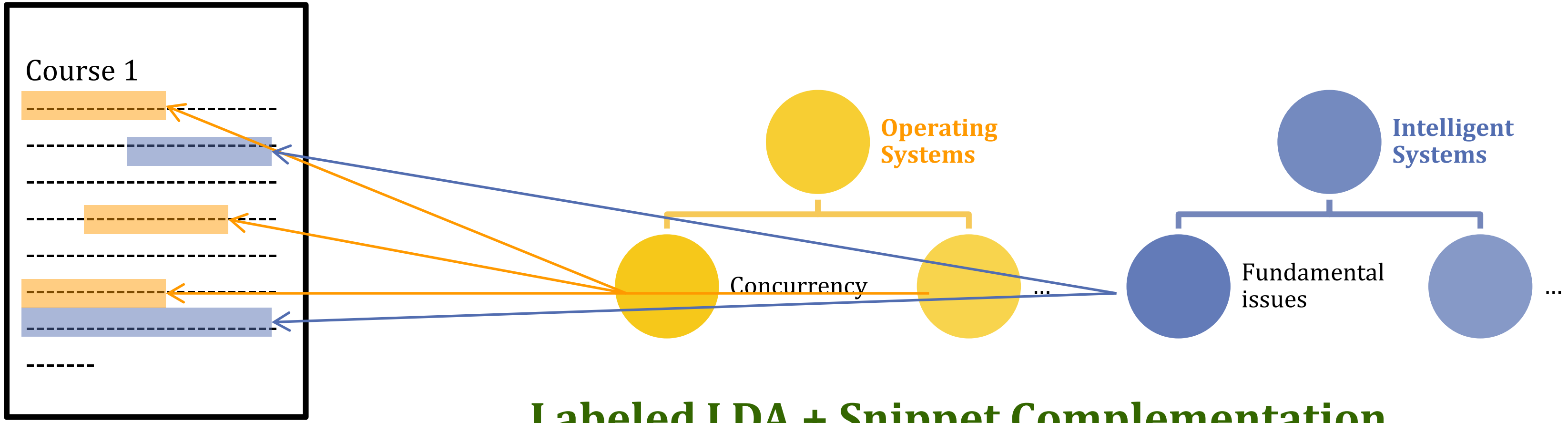
December 20, 2013

The Joint Task Force on Computing Curricula
Association for Computing Machinery (ACM)
IEEE Computer Society

A Cooperative Project of

Curriculum guideline

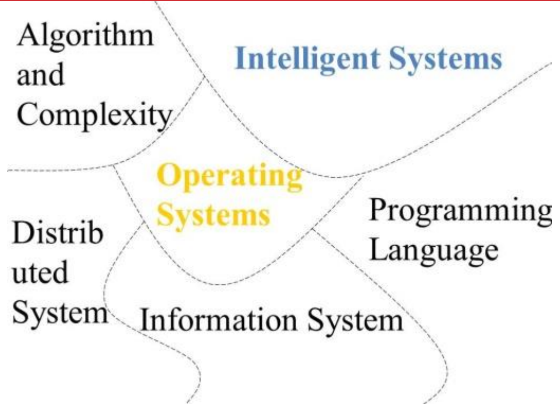
Computer Science Curricula 2013 (IEEE & ACM)



Labeled LDA + Snippet Complementation

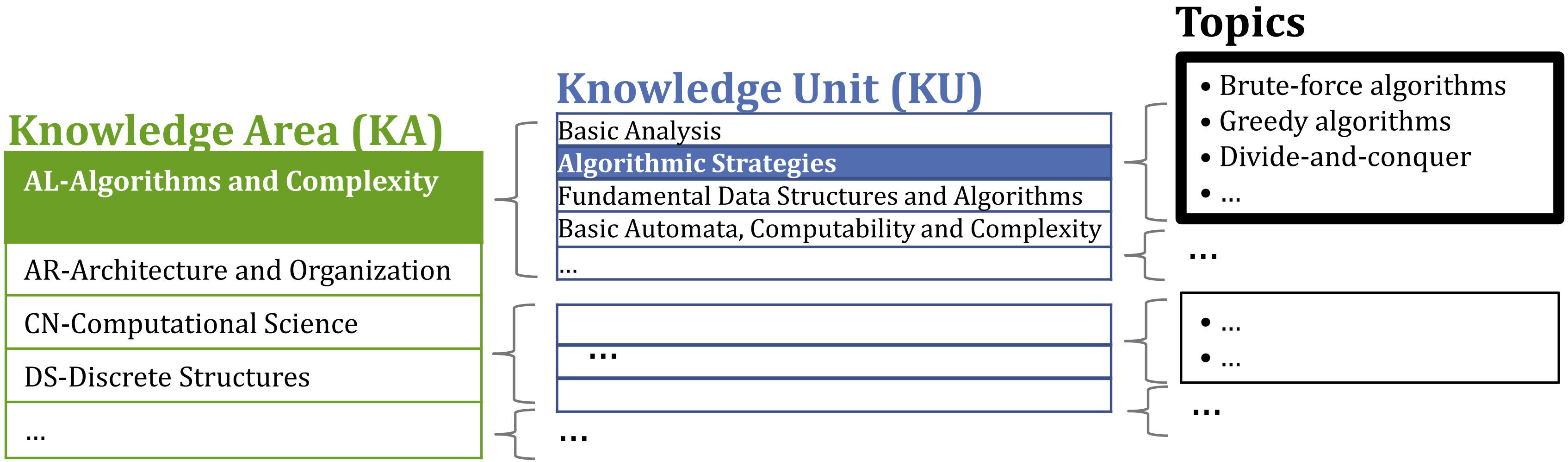
An initiative step

Course 1



Course content analysis /Dataset

CS2013: Body of Knowledge (BoK)



Topics

Knowledge Unit (KU)

Knowledge Area (KA)

AL-Algorithms and Complexity

AR-Architecture and Organization

CN-Computational Science

DS-Discrete Structures

…

Basic Analysis

Algorithmic Strategies

Fundamental Data Structures and Algorithms

Basic Automata, Computability and Complexity

…

…

…

- Brute-force algorithms
- Greedy algorithms
- Divide-and-conquer
- …

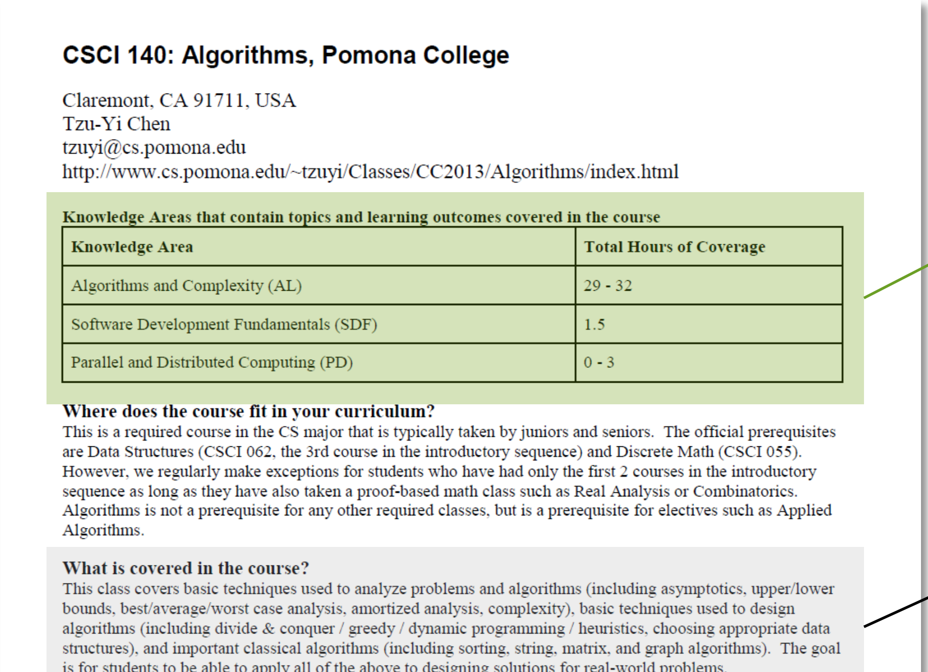
…

- …
- …

…

Course content analysis /Dataset

CS2013: Sample Courses



CSCI 140: Algorithms, Pomona College

Claremont, CA 91711, USA
Tzu-Yi Chen
tzuyi@cs.pomona.edu
http://www.cs.pomona.edu/~tzuyi/Classes/CC2013/Algorithms/index.html

Knowledge Areas that contain topics and learning outcomes covered in the course

Knowledge Area	Total Hours of Coverage
Algorithms and Complexity (AL)	29 - 32
Software Development Fundamentals (SDF)	1.5
Parallel and Distributed Computing (PD)	0 - 3

Where does the course fit in your curriculum?

This is a required course in the CS major that is typically taken by juniors and seniors. The official prerequisites are Data Structures (CSCI 062, the 3rd course in the introductory sequence) and Discrete Math (CSCI 055). However, we regularly make exceptions for students who have had only the first 2 courses in the introductory sequence as long as they have also taken a proof-based math class such as Real Analysis or Combinatorics. Algorithms is not a prerequisite for any other required classes, but is a prerequisite for electives such as Applied Algorithms.

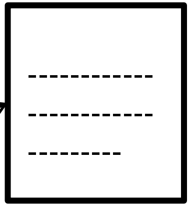
What is covered in the course?

This class covers basic techniques used to analyze problems and algorithms (including asymptotics, upper/lower bounds, best/average/worst case analysis, amortized analysis, complexity), basic techniques used to design algorithms (including divide & conquer / greedy / dynamic programming / heuristics, choosing appropriate data structures), and important classical algorithms (including sorting, string, matrix, and graph algorithms). The goal is for students to be able to apply all of the above to designing solutions for real-world problems.

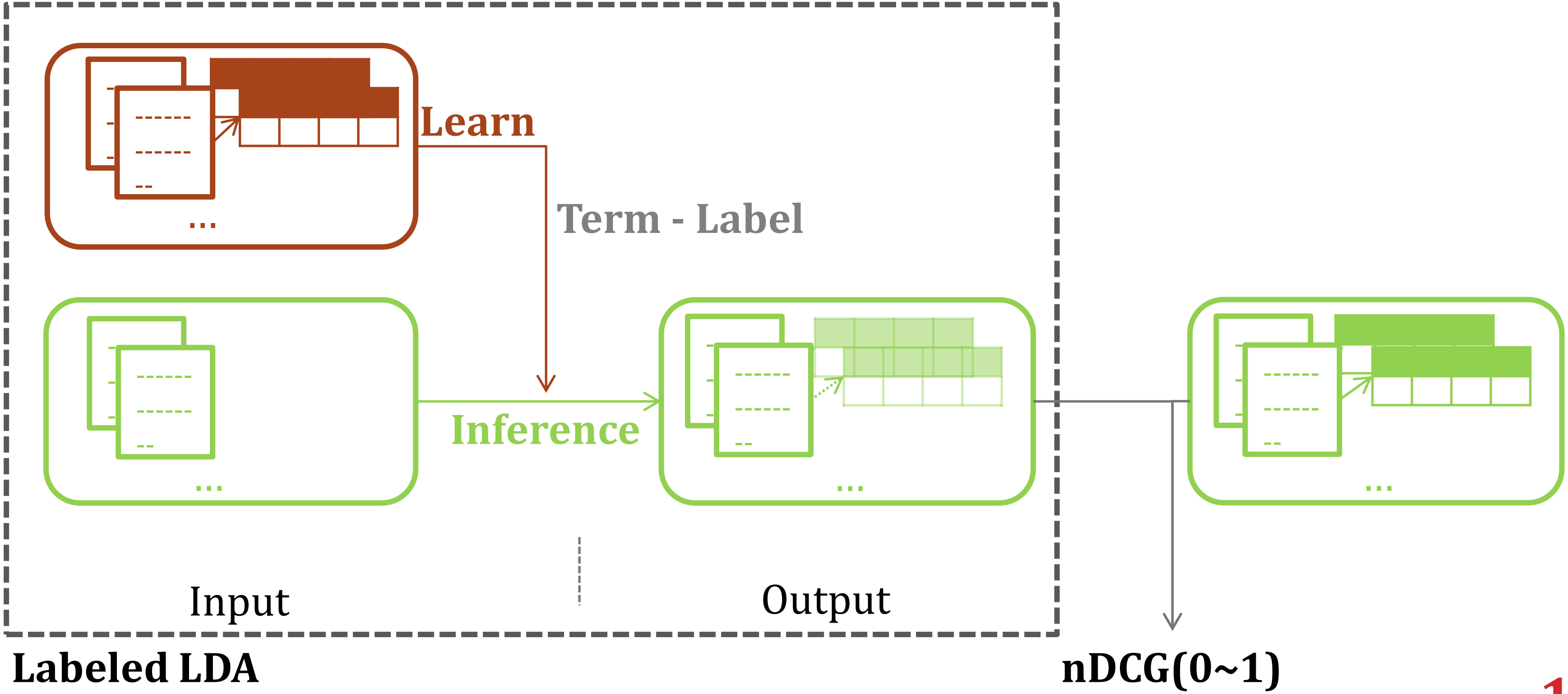
Knowledge coverage

AL	SDF	PD	...
30	1.5	1.5	...

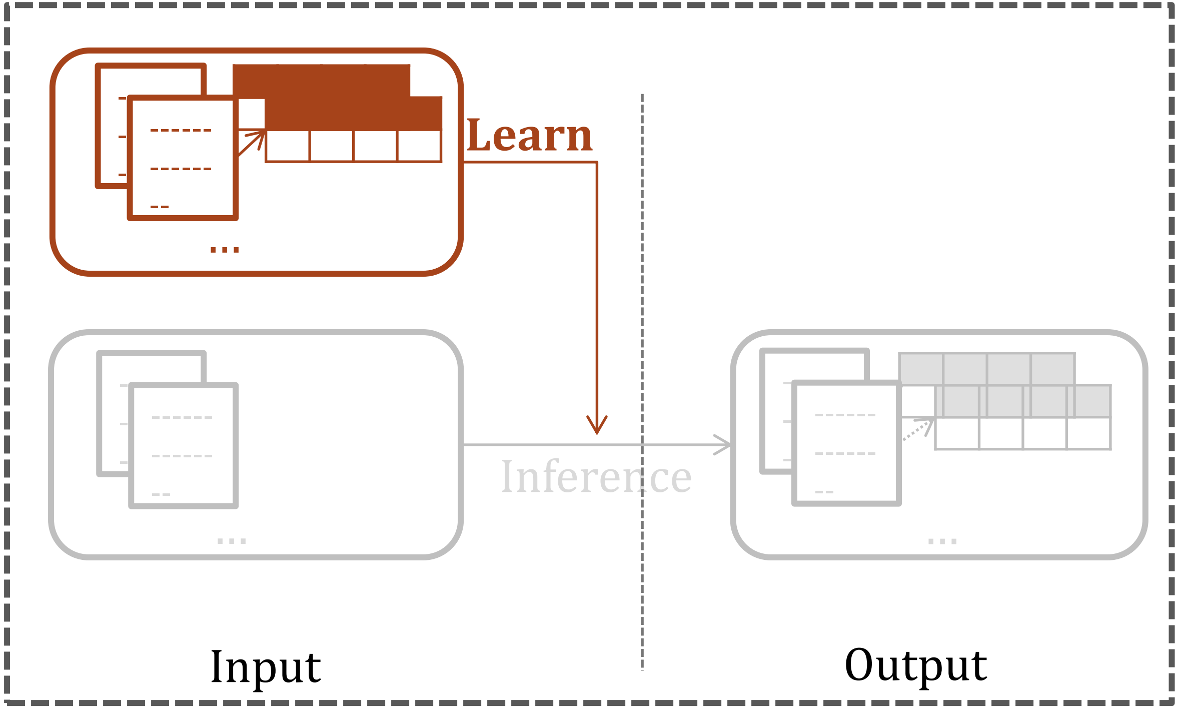
Course syllabus



Course content analysis /Method



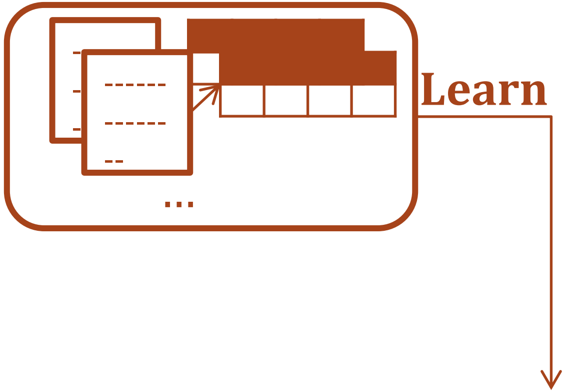
Course content analysis /Method



Labeled LDA

- **In Labeled LDA, the quality of the set of documents for learning (training set) affects the result remarkably.**

Course content analysis /Method



Documents for training:

- ■ A part of syllabi of sample courses
  - Multi-labeled
  - Combination of written language and spoken language
- ■ Body of Knowledge (CS2013)
  - Single -labeled
  - Concise language

Why? How?

Complement it with snippets

E.g.

Algorithmic strategies (AL)

- Brute-force algorithms
- Greedy algorithms
- Divide-and-conquer
- ...

The sentence is too concise to contain enough representative terms of this KU.

We integrate external texts that are explaining this KU.

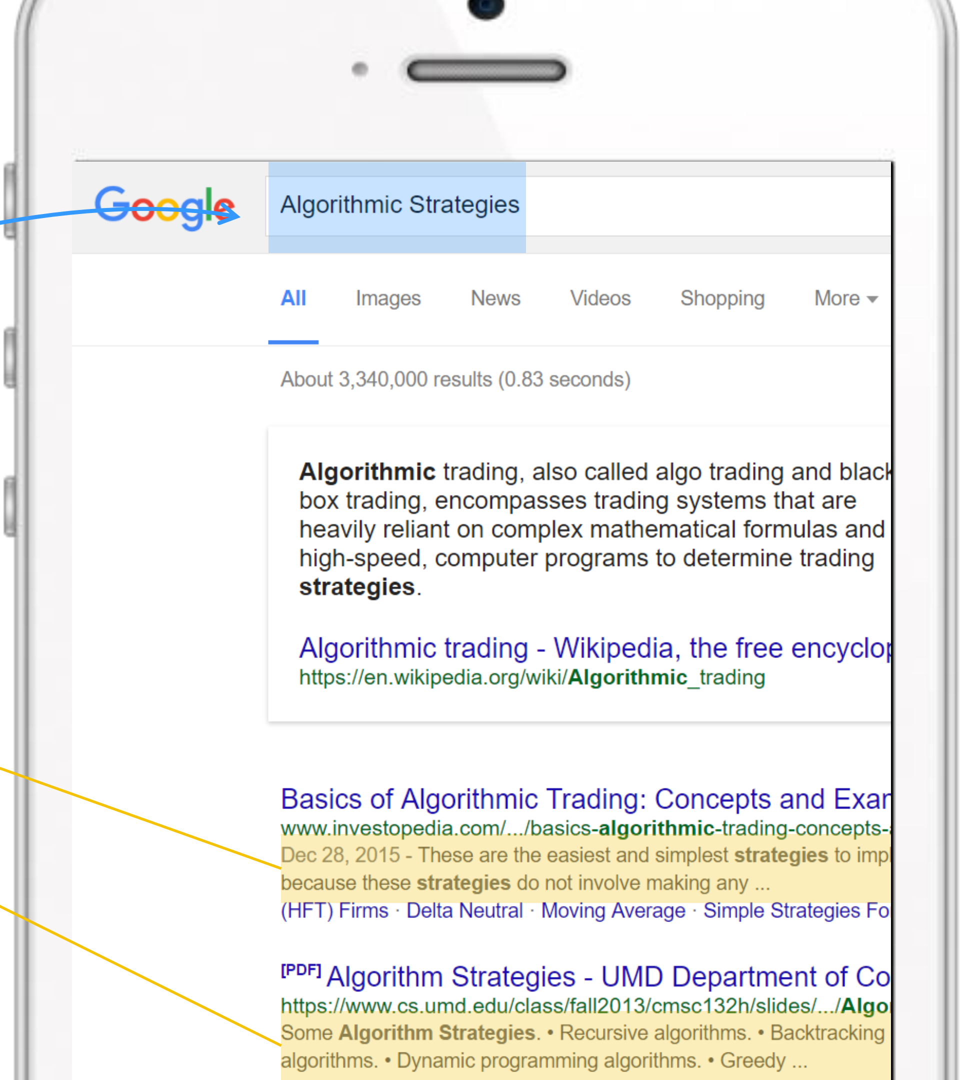
①Throw queries into search engine

Algorithmic strategies (AL)

- Brute-force algorithms
- Greedy algorithms
- Divide-and-conquer
- …

- These are the easiest and simplest…
- Some Algorithm Strategies, …

②Append snippets into the original text



About 3,340,000 results (0.83 seconds)

Algorithmic trading, also called algo trading and black box trading, encompasses trading systems that are heavily reliant on complex mathematical formulas and high-speed, computer programs to determine trading **strategies**.

Algorithmic trading - Wikipedia, the free encyclop
https://en.wikipedia.org/wiki/**Algorithmic**_trading

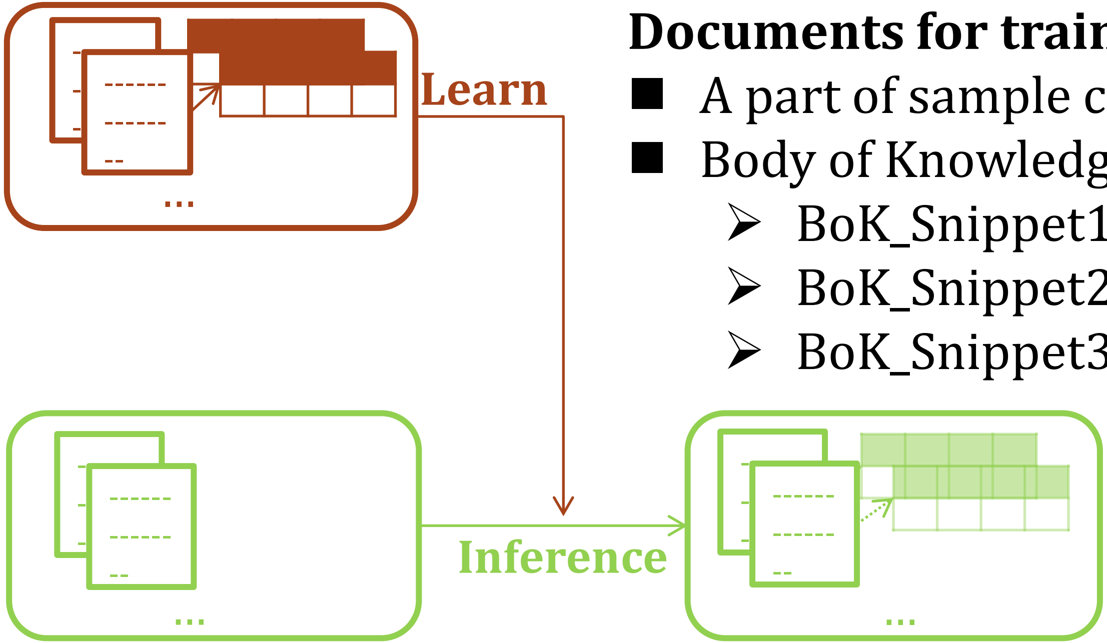
Basics of Algorithmic Trading: Concepts and Exar
www.investopedia.com/.../basics-**algorithmic**-trading-concepts-
Dec 28, 2015 - These are the easiest and simplest **strategies** to impl because these **strategies** do not involve making any ...
(HFT) Firms · Delta Neutral · Moving Average · Simple Strategies Fo

[PDF] Algorithm Strategies - UMD Department of Co
https://www.cs.umd.edu/class/fall2013/cmsc132h/slides/.../**Algo**
Some **Algorithm Strategies**. • Recursive algorithms. • Backtracking algorithms. • Dynamic programming algorithms. • Greedy ...

Patterns of queries to extract snippets:

Query	Result
The title of the KU	snippet1
The title of the KU and the title of the KA it belongs to	snippet2
Top three representative terms of the KU (tf-idf value)	snippet3

Course content analysis /Method

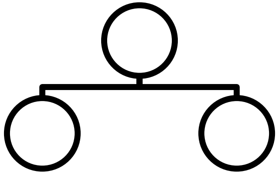


Documents for training:

- A part of sample courses (SC)
- Body of Knowledge (BoK)
  - BoK_Snippet1(KU title)
  - BoK_Snippet2 (KU+KA title)
  - BoK_Snippet3 (KU top terms)

Labels to inference:

- KA
- KU



Course content analysis /Method

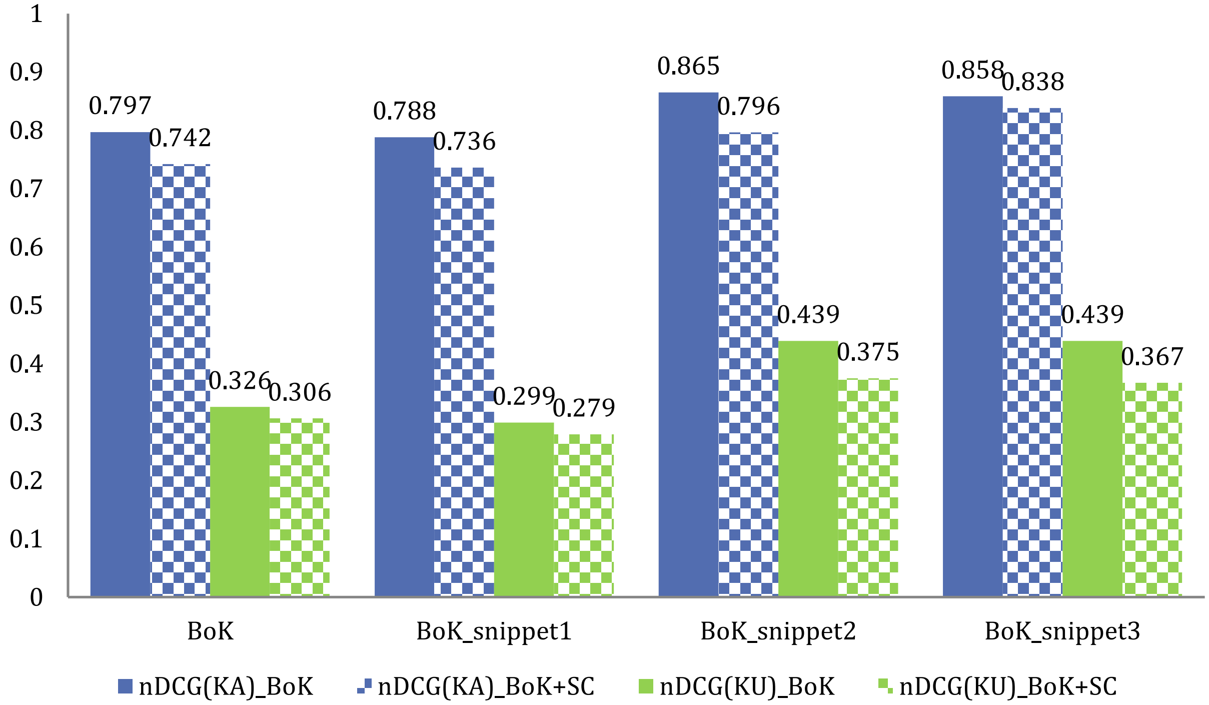
Experiments:

	No snippet	Snippet1	Snippet2	Snippet3
BoK	BoK	BoK_snippet1	BoK_snippet2	BoK_snippet3
BoK+SC	BoK+SC	BoK_snippet1+SC	BoK_snippet2+SC	BoK_snippet3+SC

We want to observe:

- Which type of documents is suitable for training?
- Does snippet work?
- Does this model work for both KA and KU?

Course content analysis /Result



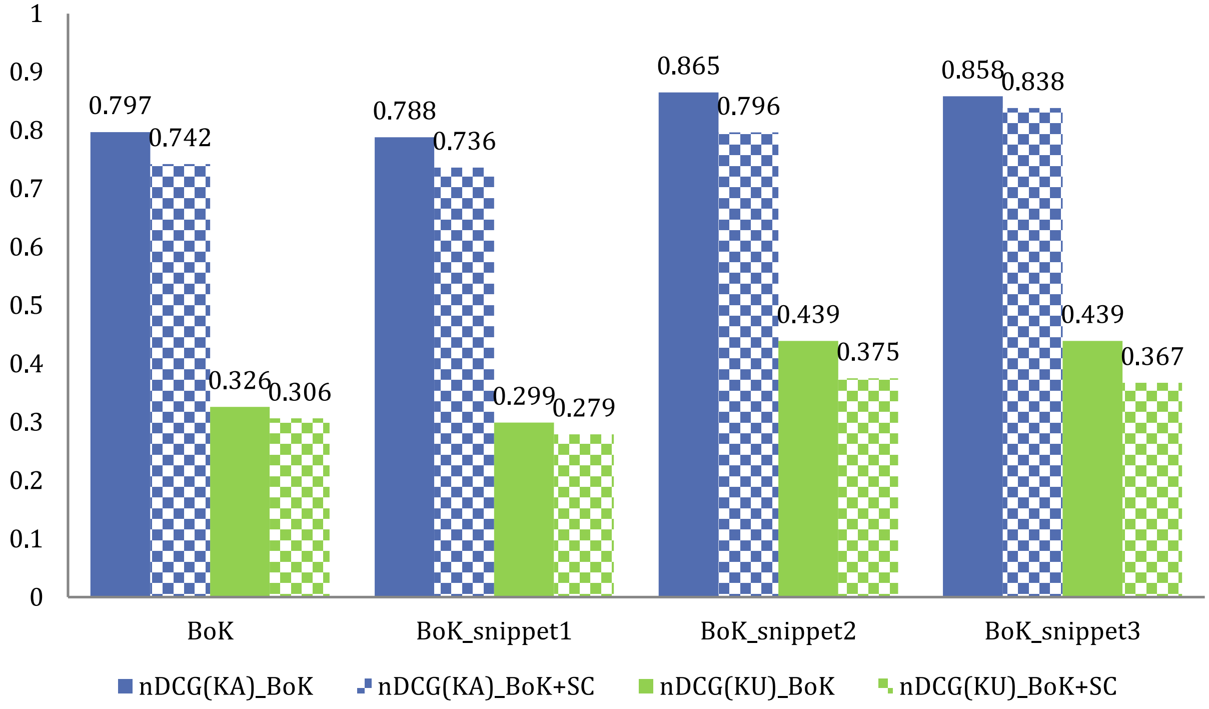
Result 1:

BoK > BoK + SC

Discussion 1:

The documents of SC are multi-labeled, which brings some noise in Labeled LDA.

Course content analysis /Result



Result 2:

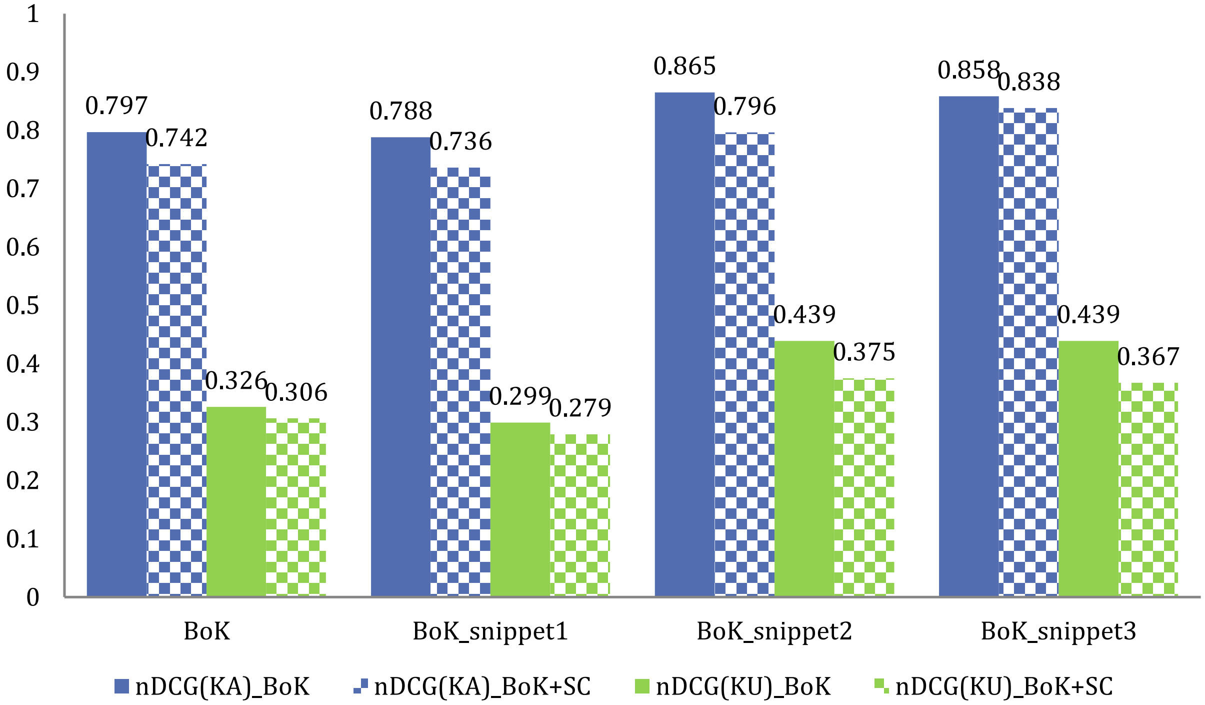
Snippet2, Snippet3☑

(KU+KA title, top 3 terms)

Discussion 2:

We can complement an insufficient domain knowledge base by information retrieval techniques

Course content analysis /Result



Result 3:

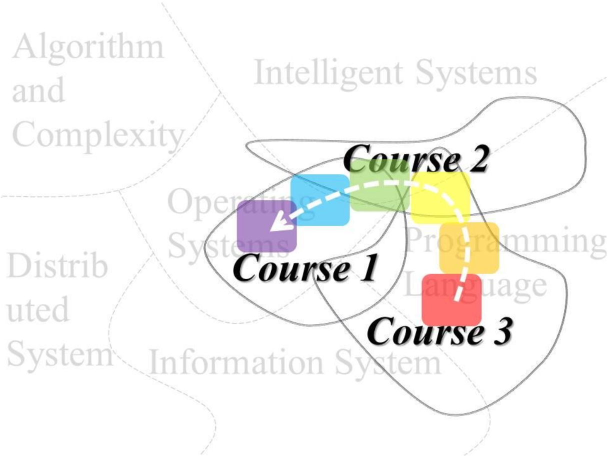
KA☑ KU?

Discussion 3:

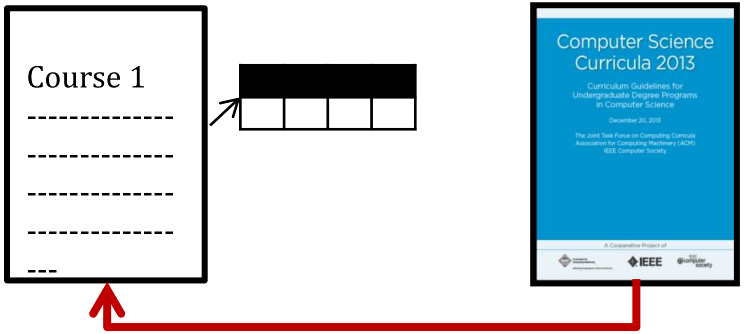
It needs to be improved when extracting knowledge coverage at a smaller granularity.

SUMMARY

Recommendation system design



Course content analysis


Course 1

Computer Science Curricula 2013
Curriculum Guidelines for Undergraduate Degree Programs in Computer Science
December 20, 2013
The Joint Task Force on Computing Curricula
Association for Computing Machinery (ACM)
IEEE Computer Society
A Cooperative Project of
IEEE

Labeled LDA

+ Snippet Complementation

Q&A

Yiling Dai (daiyiling10@hotmail.com)