Automating Gap Analysis of Learning Outcomes

Vijay Mago, Department of Computer Science, Lakehead University, Thunder Bay
April 16, 2018
Automating Gap Analysis of Learning Outcomes through Natural Language Processing
2017-17-ONCAT
Text Extraction

Natural Language Processing

Visualization of results
Extracting Learning Outcomes

Introduction to Databases

Syllabus

Web Page
http://www.cs.northwestern.edu/~pdinda/db

Instructor
Peter A. Dinda
1890 Maple Avenue, Room 338
847-467-7859
pdinda@cs.northwestern.edu
Office hours: Thursdays, 2-4pm or by appointment

Teaching assistants
Ananth Sundararaj
1890 Maple Avenue, Room 332
847-491-7159
ais@cs.northwestern.edu
Office hours: Mondays, 10:15am-12:15pm, Wednesdays 11:15am-12:15pm
or by appointment

Bin Lin
1890 Maple Avenue, Room 224
847-491-7159
binline@cs.northwestern.edu
Office hours: Tuesdays, 10-12am, Wednesdays 3:30-4:30pm
or by appointment

Location and Time
1890 Maple Avenue, CS Department classroom, MWF 9-9:50am

Prerequisites
Required
CS 311 or equivalent data structures course

Highly recommended
CS 213 or equivalent computer systems course
Familiarity with concepts from discrete math such as set theory
Some familiarity with Perl or other scripting language

COURSE OUTLINE

SCHOOL: School of Engineering Technology and Applied Science
DEPARTMENT: Information and Communication Engineering Technology (ICET)
PROGRAM: Software Engineering Technician/Technologist (3109, 3119, 3409, 3419, 3508, and 3508)

COURSE TITLE: Introduction to Databases
COURSE CODE: COMP122

TOTAL COURSE HOURS: 60 Hours
PRE-REQUISITES/CO-REQUISITES: COMP100

COURSE ELIGIBILITY FOR PRIOR LEARNING ASSESSMENT AND RECOGNITION (PLAR): Yes

ORIGINATED BY: Bhim Harlal
REVISED BY: Bhim Harlal
DATE: August 2009

APPROVED BY:
Chairperson/Dean

Automating Gap Analysis of Learning Outcomes

ONCAT 2018
Natural Language Processing

- Dependency parsing to establish syntactical relationship between words

- Example 1: *Prepare financial statements and related disclosures*

Fig 1. Dependency parsing
Natural Language Processing

- Dependency parsing to establish syntactical relationship between words

- Example 2: *Analyze and account for complex business transactions*

Fig 2. Dependency parsing

Both the learning objectives are parsed separately and the parsing information is compared.
Data Extraction - Challenges

Course outlines do not have a defined format

Assessing the relevance of the text in the document

Documents may have header/footer text

Text might be divided into text columns, which makes it important to analyze the layout of the document before proceeding with text extraction.
Natural Language Processing

1. Similarity between peculiar words in a domain
e.g. in computer science, we have different programming languages: Java, Python, C, C++. Using a general purpose corpus (vocabulary) would not yield precise similarities between such terms.

2. Size of the corpus pertaining to a specific domain
We decided to use Wikipedia as corpus since it covers all the domains and is constantly updated with the new terms and content.

3. Cleaning the textual content from Wikipedia pages to get rid of ASCII characters, URLs and unnecessary tags.

4. Forming collocations in the entire corpus and then training the model
e.g. replacing “computer science” with “computer_science” and “programming languages” with “programming_languages”

5. An algorithm to establish similarity between two learning objectives and extending it to cover broader scope such as course comparison and program comparison
Website: http://www.loaga.science

Test User Credentials:
   Username: testuser1@loaga.science
   Password: password

Contact:
   Vijay Mago: vmago@lakeheadu.ca
Bloom’s Taxonomy

Though Bloom’s taxonomy is the suggested standard for designing the course outline, we have found that a considerable number of course drafts differ significantly from the norm. To use the Bloom’s taxonomy, we establish the 'Bloom Index'. The Bloom Index represents the gap between two learning outcomes according to the verbs in LOs.
Finding Learning Outcome

1. Extracting Text Format Information (3.1)
2. Detecting Relevant Headings (3.2)
3. Selecting Beginning and End Marker (3.3)
4. Layout Analysis Using White Space Analysis (3.4)
5. Targeted Text Extraction (3.5)
6. Formatting Output (3.6)
Visualizing the Mapping of Outcomes, Content, and Curriculum between programs to support transfer

Nerissa Mulligan, Brian Frank, Roderick Turner, Mary Pierce, Jake Kaupp, Vijay Mago
2015 ONCAT Project: Framework

Math
- Integral and differential calculus

Natural and Engineering Science
- Mechanics and Electro-magnetism

Design & Investigation
- Engineering design
- Team dynamics
- Project management
- Communication
- Lab skills

2016 ONCAT Project: Comparisons

Engineering Programs
- Electrical Engineering
- Mechanical Engineering

Technology Programs
- Electrical / Electronic Engineering
- Mechanical Engineering Technology
Content vs. complexity

First year calculus content
Functions, limits, derivatives; optimization, rate problems, exponentials, logarithms, inverse trigonometric functions; exponential growth as an example of a differential equation. Fundamental Theorem of Calculus, Riemann integral; applications to problems involving areas, volumes, mass, charge, work, etc. Some integration techniques.

Cognitive complexity?
Novelty of problems?
How scaffolded?
Learning Outcomes

- **analyze**, **interpret**, and **produce** electrical and electronics **drawings**, **technical reports** including other related documents and graphics.

**Concept:** engineering drawing

**Bloom’s cognitive level:**
- analyze
- interpret
- produce
Anderson and Krathwohl, 2001: Cognitive Process
SOLO Taxonomy: Structural complexity
## Categorizing Instructor’s Scaffolding

<table>
<thead>
<tr>
<th>Prescribed</th>
<th>The activity instructs the student to follow a prescribed sequence of calculations or an explicitly stated approach.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained</td>
<td>The activity description does not specify the solution; the general approach is implied through question sequencing, headings, etc.</td>
</tr>
<tr>
<td>Scaffolded</td>
<td>The activity requires the student to choose from a range of approaches.</td>
</tr>
<tr>
<td>Adopted</td>
<td>The activity requires the student to synthesize different methods and formulate novel methods or apply existing ones to novel applications.</td>
</tr>
</tbody>
</table>
Learning outcomes assessments for two core courses in Electrical/Electronic engineering in (a) technology and (b) engineering exam questions as scored by content specialists using a three-part framework.
Goal: Automate some analysis and visualization
Focus Groups

• Two focus groups

• 19 Institutions

• 31 Participants from 6 different programs

Academic Quality Lead
Academic Manager (Nursing Program)
Admission Assistant
Assistant Professor
Assistant Registrar
Associate Dean
Associate Professor
Business Program Contact
Chair (Admissions Committee)
Credit Transfer Officer
Curriculum Consultant
Dean

Director (School of Kinesiology)
Director (Centre for Academic Excellence)
Enrolment Services and Strategic Partnerships
International Recruitment Officer
Registrar’s Office
Manager
Transfer Credit Advisor
Pathways and Credit Transfer Coordinator

Program Coordinator
(Bachelor of Applied Health Information Science)
Program Head
Program Manager
Student Advisor
Undergraduate Academic Advisor
1. What information do you typically have when assessing student transfers coming into (or going out of) your program?

2. What information would you ideally like to have (but do not) when assessing student transfers coming into (or going out of) your program?

   - transcripts
   - course syllabi
   - CLO
   - PLO
   - institutional information
   - course success
   - course assessment tools
   - course textbook lists
   - other

3. What do you think are the most common barriers to transfers in your program?

4. Please tell us a bit about any tools, assessments or methodologies that you have found successful when dealing with transfers that you can share with us.
## Transfer Information

<table>
<thead>
<tr>
<th>Program</th>
<th>Transcripts</th>
<th>Course syllabi</th>
<th>CLOs</th>
<th>PLOs</th>
<th>Institutional information</th>
<th>Course success rates</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>⊙</td>
</tr>
<tr>
<td>5</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>9</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>12</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>♦</td>
</tr>
<tr>
<td>17</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Transfer Information Wish List

<table>
<thead>
<tr>
<th>Program</th>
<th>Transcripts</th>
<th>Course syllabi</th>
<th>CLOs</th>
<th>PLOs</th>
<th>Institutional information</th>
<th>Course success rates</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td>☄</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>9</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>☯</td>
</tr>
<tr>
<td>17</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>☯</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Primary Barriers to Transfer

Time
Cost
Equivalency
Changes to Courses
Availability of Information
Course Alignment
Accreditation
Age of Credits
Primary Barriers to Transfer

- Time
- Cost
- Equivalency
- Changes to Courses

Availability of Information

Course Alignment

- Accreditation
- Age of Credits
Tips for assessing transfer

- Course syllabi online
- Pathways
- Bridging Programs
- Students
- Learning Outcomes
- Percentage overlap
- Database
- Shared folders
- Official transfer credits
- Backwards design & course-to-course
- Credit Transfer Evaluation Guide
- Transfer agreements for business
Focus Group Questions

1. What information do you use to evaluate transfer now?

- Title
- Description
- Learning Outcomes
- Assessment Criteria
- Textbook list
- Degree of difficulty
- Work samples
- Program calendar
- Credit hours
- Program accreditation map
Focus Group Questions

2. If you were to setup a multi-institutional transfer agreement between diploma programs and degrees in your discipline, what information would you use?

- Past performance of students
- Well-written learning outcomes
- Context of learning outcomes
- Conversations between instructors
- Accreditation
- Professional registration status of instructor
Focus Group Questions

3. What analysis, comparison, and/or visualization would you find useful to accomplish Q2?

- Course-to-course
- Program-to-program
- Gap Analysis
- Heat map
- Beyond Bloom’s
- Institutional matches
- Record of past equivalencies
- Ability to “drill down”
create visualizations to help support student transfer and the creation of pathways in Ontario.
Web Application Requirements

- Easy upload of course materials
- Automated categorization of outcome verb, curriculum structure, prerequisites.
- Simple user interface
- Output visualizations
- Draw on existing tools when possible.
- Ability to save imported programs
- Code publically available.
Original Functional Requirements

Input:
• Pull information in real-time
• Access PLOs, CLOs, WLOs
• Textbook lists
• Transcripts
• Adaptable

Output:
• Course-level overlap analysis
• Gap analysis
• Information about existing pathways
• Customizable
• Flexible
Web Application Elements

Information Elements

• Landing page (login)
• FAQ
• How to Use
• Background
• Contact Us

Tool Elements

User uploads or inputs syllabus OR search and filter course curricula by institution, year, curricula code

Parse outcomes and course name

Classify outcome verb using Bloom’s and SOLO taxonomy

Make visualizations

Export and save results
An Interactive app that compares multiple course outlines semantically

Start Comparing

How to use

Upload Document

Compare Learning Outcomes

Visualize Learning Outcomes

Analyze The Data
Create an Account

First Name

Last Name

Email

Password

Sign up
Program Details

Institute

Program Name

Course Document

Drag & Drop a File

or

Upload

Process
Program Details

Institute

Program Name

Course Document

Drag & Drop a File

or

Upload

MECH 319 - Fabrication Design - Fall 2016

Course Name

Quality Control 1

Learning Outcomes

Describe the capability of machinery for sheet metal component manufacturing 1
<table>
<thead>
<tr>
<th></th>
<th>Remember</th>
<th>Understand</th>
<th>Apply</th>
<th>Analyze</th>
<th>Evaluate</th>
<th>Create</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extended Abstract</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relational</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi-structural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unistructural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-structural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Engineering**

**Engineering technology**
Learning Objective Automated Gap Analysis

Vijay Mago
# Progress

<table>
<thead>
<tr>
<th>Task</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application Design</td>
<td>Completed</td>
</tr>
<tr>
<td>Solo/bloom tree map</td>
<td>Completed</td>
</tr>
<tr>
<td>Solo/Bloom heatmap</td>
<td>Completed</td>
</tr>
<tr>
<td>Pie chart(Additional)</td>
<td>Completed</td>
</tr>
<tr>
<td>Export visuals as pdf/png format</td>
<td>Completed</td>
</tr>
<tr>
<td>Add new course to existing program</td>
<td>Completed</td>
</tr>
<tr>
<td>Add more than one program in one session</td>
<td>Completed</td>
</tr>
<tr>
<td>Store user’s sessions and comparisons</td>
<td>In progress</td>
</tr>
<tr>
<td>Show details of courses user uploads</td>
<td>In progress</td>
</tr>
<tr>
<td>Share by email</td>
<td>In progress</td>
</tr>
<tr>
<td>Design of Other pages</td>
<td>In progress</td>
</tr>
<tr>
<td>Extraction of learning outcomes and course name from pdf</td>
<td>In progress</td>
</tr>
</tbody>
</table>
Activity: Small groups

Review the visualizations

1. Do you think these current visualizations provide useful information? How could they be used?
2. How could the visualizations be adapted, or new analysis added, to support building transfer pathways?

PollEv.com/brianfrank116
Where this is going

Develop proposal for province-wide pathway between engineering technology and engineering

Continue developing the app to support pathway development

Identify other groups who would like to adapt and expand the tool
Visualizing the Mapping of Outcomes, Content, and Curriculum between programs to support transfer

Nerissa Mulligan, Brian Frank, Roderick Turner, Mary Pierce, Jake Kaupp, Vijay Mago

Brian.Frank@QueensU.ca, @BFrankQueensu
Roderick.Turner@senecacollege.ca
MPierce@fanshawec.ca