A Novel Approach for Test Problem Assessment Using Course Ontology

Manas Hardas

Department of Computer Science
Kent State University, OH.

Oct 23rd 2006
Objective

Design is useful for reengineering and reuse

Basic unit of test, is a problem

To answer a problem, knowledge is required
Background

- Web is scattered with online educational resources

- Mostly un-organised, but some in organised fashion as well [OCW, Universia, ACM, NSDL, CORE]

- **Not represented in context**

- Looses reusability, reengineering not possible, not machine interpretable

- Semantic representation standards
  - RDF ([http://www.w3.org/RDF/](http://www.w3.org/RDF/)) (2002)
  - OWL ([http://www.w3.org/TR/owl-features/](http://www.w3.org/TR/owl-features/)) (2004)

- **Contextual representation of problems is important**
Resources represented in knowledge context using RDF/OWL

Knowledge Domains

Assess Problems for knowledge

Test Design

Test Problems not in context

WEB
In an Ethernet network can a second packet be transmitted as soon as the receiver receives the first packet?

Objectives

• map concept knowledge
• assess test problems
• analyze the methodology
Scope of this talk

• Course Knowledge Representation
• Problem Assessment
• Results
CSG (concept space graph)

- Course Knowledge is represented using Concept Space Graph called as a “Course Ontology”

- Course ontology → hierarchical representation of concepts taught in a course linked by “has-prerequisite” relationships.

- Each link → has prerequisite, link weight

- Each node → Self-weight, prerequisite-weight

- Expressive

- Computable

10/25/2006
Course Ontology Description Language (CODL)

• Written in Web Ontology Language (OWL)
• Mostly OWL Lite with few extensions on data type properties
• Can represent any course ontology
• Basic Elements on Course Ontology OWL document are
  • Ontology Header
  • Class descriptions
  • Property descriptions
  • individuals
CODL individuals

Individuals

<Concept rdf:ID="MemoryManagement"/>

<Concept rdf:ID="OS">
    <hasPrerequisite>
        <Relation rdf:ID="relation_1">
            <connectsTo rdf:resource="#MemoryManagement"/>
            <hasLinkWeight rdf:resource="#0.2"/>
        </Relation>
    </hasPrerequisite>
    <hasSelfWeight rdf:resource="0.39"/>
    <hasPrerequisiteWeight rdf:resource="0.61"/>
</Concept>
Problem Assessment Methodology
Problem – Concept Mapping
\[C_1, C_2, \ldots, C_n\]
Threshold coefficient \(\lambda\)

Course ontology

Problem Assessment System

CSG Extraction Module
Projection Calculation Algorithms

Assessment Module
Evaluation Parameters Calculation Algorithms

Assessment Parameters
\([\alpha, \Delta, \delta]\)

Threshold coefficient \(\lambda\)

Problem – Concept Mapping
\[C_1, C_2, \ldots, C_n\]

Threshold coefficient \(\lambda\)

Course ontology

Problem Assessment System

CSG Extraction Module
Projection Calculation Algorithms

Assessment Module
Evaluation Parameters Calculation Algorithms

Assessment Parameters
\([\alpha, \Delta, \delta]\)
CSG extraction

Why?
- CSG is very big
  - WordNet 50,000 word
  - CYC (over a million assertions)
  - Medical/Clinical Ontology (LinKBase 1 million concepts)
- Selection of relevant portion of ontology to maintain computability

How?
- Projection Graph
- Projection Threshold Coefficient ($\lambda$)
  - Prunes CSG
  - Desired semantic depth
How is routing achieved in Autonomous Systems?
Prerequisite effect of one node over another

• **Node Path Weight:** When two concepts \(x_0\) and \(x_t\) are connected through a path \(p\) consisting of nodes given by the set \([x_0, x_1, \ldots, x_m, x_{m+1}, \ldots, x_t]\) then the node path weight between these two nodes is given by:

\[
\eta(x_0, x_t) = W_s(x_t) \prod_{m=t}^{1} [l(x_{m-1}, x_m) * W_p(x_{m-1})]
\]

The node path weight for a node to itself is its self weight: \(\eta(x_1, x_1) = W_s(x_1)\)

• **Incident Path Weight:** It is the “the absolute prerequisite cost required to reach the root node from a subject node.” Incident path weight is same as node path weight without the factor of self weight of the subject node.

\[
\gamma(x_0, x_n) = \frac{\eta(x_0, x_n)}{W_s(x_n)} = \frac{\eta(x_0, x_n)}{\eta(x_n, x_n)}
\]
Example CSG(A)
Node path weight, Incident Path Weight calculations

$$\eta(B, L) = W_s(L) \prod [l(F, L) \cdot W_p(F) \cdot l(B, F) \cdot W_p(B)]$$

$$\eta(B, L) = 0.3 \cdot 0.5 \cdot 0.8 \cdot 0.55 \cdot 0.8 = 0.0528$$

$$\gamma(B, L) = \eta(B, L) / W_s = 0.0528 / 0.3 = 0.176$$

$$\eta(B, L) = W_s(L) \prod [l(E, L) \cdot W_p(E) \cdot l(B, E) \cdot W_p(B)]$$

$$\eta(B, L) = 0.3 \cdot 0.15 \cdot 0.6 \cdot 0.4 \cdot 0.8 = 0.00864$$

$$\gamma(B, L) = \eta(B, L) / W_s = 0.00864 / 0.3 = 0.0288$$
Projection graph

- Given a root concept $x_0$ and a projection threshold coefficient $\lambda$, and CSG, $T(C, L)$, a projection graph $P(x_0, \lambda)$ is defined as a sub graph of $T$ with root $x_0$ and all nodes $x_t$ where there is at least one path from $x_0$ to $x_t$ in $T$ such that node path weights satisfies the condition: 
  $$\eta(x_0, x_t) \geq \lambda$$

The projection set consisting of nodes $[x_0, x_1, x_2, \ldots, x_n]$ for a root concept $x_0$ is represented as, 
  $$P(x_0, \lambda) = P^{x_0} = [x_0^{x_0}, x_1^{x_0}, x_2^{x_0}, \ldots, x_n^{x_0}]$$

Where $x_i^j$ represents the $j^{th}$ element of the projection set of node $j$. 
Projection Calculation Example

Calculate the projection graph for Concept B, for $\lambda=0.001$. 

\[
\begin{array}{c|c}
\text{Node} & \eta(r,n) \\
\hline
B & 0.0024 \\
E & 0.0026 \\
F & 0.0036 \\
G & 0.0012 \\
I & 0.0012 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\text{Local root} & \text{Node} & \text{Local root} & \eta(r,n) \\
\hline
B & E & \checkmark & \checkmark \\
F & F & \checkmark & \checkmark \\
I & C & \checkmark & \checkmark \\
\end{array}
\]
### Projection Calculation

#### Node "n" Projection Calculation

<table>
<thead>
<tr>
<th>Local root “r”</th>
<th>Node “n”</th>
<th>$\eta(r,n)$</th>
<th>$\eta(r,n) \geq \lambda$?</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>G</td>
<td>0.00125</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>0.02125</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>0.005</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0.00357</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>0.00475</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.00028</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>0.00034 (H)</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.00675 (I)</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>0.0135</td>
<td>✓</td>
</tr>
</tbody>
</table>

#### Diagram

![Diagram](image-url)
Problem Assessment Parameters
Coverage

- Knowledge required

Coverage of a node $x_0$ with respect to the root node $r$ is defined as the product of the sum of the node path weights of all nodes in the projection set $P(x_0, \lambda)$ for the concept $x_0$ and the self weight of $x_0$ and the incident projection path weight $\gamma (r, x_0)$ from the root $r$.

If the projection set for concept node $x_0$, $P(x_0, \lambda)$ is given by $[x_0, x_1, x_2...x_n]$ then the coverage for node $x_0$ about the root $r$ is defined as,

$$\alpha(x_0) = \gamma(r, x_0) \times \sum_{m=0}^{n} \eta(x_0, x_m)$$

Total coverage of multiple concepts in a problem given by set $[c_0, c_1, c_2...c_n]$ is,

$$\alpha(T) = \alpha(C_1) + \alpha(C_2) + ... + \alpha(C_n)$$
A question connects to concepts B and D from the ontology. Find its coverage.

\[ \alpha(B) = \gamma(A, B) \times \sum \eta(B, P_i^B) \]

\[ \alpha(B) = (0.5 \times 0.98) \times (0.335882) \]

\[ \alpha(B) = 0.16458218 \]
\[ \alpha(D) = \gamma(A, D) \times \sum_i \eta(D, P_i^D) \]

\[ \alpha(D) = (0.2 \times 0.98) \times (0.0560625) \]

\[ \alpha(D) = 0.01098825 \]

\[ \alpha(\text{total}) = \alpha(B) + \alpha(D) \]

\[ \alpha(\text{total}) = 0.16458218 + 0.01098825 \]

\[ \alpha(\text{total}) = 0.17557043 \]
Diversity

- The breadth of knowledge domain
- Opposite of similarity

The ratio of summation of node path weights of all nodes in the non-overlapping set to their respective roots, and the sum of the summation of node path weights of all nodes in the overlap set and summation of node path weights of all nodes in the non-overlap set.

Diversity,

$$\Delta = \frac{\sum_{m=1}^{p} \eta(i, N_m)}{\sum_{m=1}^{q} \eta(j, O_m^j) + \sum_{m=1}^{p} \eta(i, N_m)}$$

where  $$\forall i, j \in C$$

Where, Concept set,  $$C = [C_0, C_1, C_2...C_n]$$
Projection sets,  $$P(C_0, \lambda) = [x_1^C_0, x_2^C_0, ..., x_a^C_0]$$,  $$P(C_1, \lambda) = [x_1^{C_1}, x_2^{C_1}, ..., x_b^{C_1}]$$  ...  $$P(C_n, \lambda) = [x_1^{C_n}, x_2^{C_n}, ..., x_c^{C_n}]$$
Overlap set,  $$O = [O_0, O_1, O_2...O_q]^j$$
Non-overlap set,  $$N = [N_0, N_1, N_2...N_p]^j$$
Diversity Calculation

$O = [I, M, N, O, P]$

$\Delta = \frac{\sum \eta(n, N^e)}{\sum \eta(n, O^e) + \sum \eta(n, N^e)}$

$n = B \mid D; \quad x^e = \text{element of set } x$

$\Delta = \frac{1.44714}{0.0448045 + 1.44714} = 0.97$
Conceptual Distance

- **Measures similarity between concepts i.e. distance from ontology root.**

  - *It is defined as the log of inverse of the minimum value of incident path weight (maximum value of threshold coefficient) which is required to encompass all the concepts from the root concept.*

If question asks concept set \( C = [C_0, C_1, C_2...C_n] \) then the conceptual distance from the root concept \( r \) is given by,

\[
\delta(C_0, C_1...C_n) = \log_2 \left( \frac{1}{\min[\gamma(r, C_0), \gamma(r, C_1)...\gamma(r, C_n)]} \right)
\]

- Greater the distance between the concepts, more is the semantic depth.
Conceptual Distance Calculation

Calculate conceptual distance between (E, F, M)

\[
\delta(E,F,M) = \log_2 \left( \frac{1}{\min\{\phi(A,E), \phi(A,F), \phi(A,M)\}} \right)
\]

\[
\delta(E,F,M) = \log_2 \left( \frac{1}{\min\{0.1568, 0.2156, 0.002375\}} \right)
\]

\[
\delta(E,F,M) = \log_2 (429.65)
\]

\[
\delta(E,F,M) = 2.63
\]
Results and Parameter Performance Analysis
Setting

- Operating system course ontology created using prescribed text books
- OSOnto (>1350 concepts)
- XML and OWL

4 quizzes, 38 questions composed using concepts selected from OS Ontology

Tests administered by at least 25 graduate and undergraduate students

Scoring done by at least 2 graders per question and average score taken.

Do the parameters provide any insight into the perceived difficulty/complexity of the question?

Performance analysis = Plotting average score/parameter values
coverage vs. average score

- coverage and average score inversely correlated
- behavior constant for changing threshold coefficient
Diversity vs. average score

- diversity and average score inversely correlated
- behavior constant for changing threshold coefficient
Conceptual distance vs. average score

- Conceptual distance and average score inversely correlated
- Distance does not vary with threshold coefficient
Correlation study

- coverage-avg. score correlation decreases with threshold coefficient
- diversity-avg. score correlation decreases with threshold coefficient
Observations and Inferences

- (Coverage, diversity and conceptual distance) $\alpha \ (1/\text{Average score})$
  - Indicates perceived difficulty
  - Coverage gives the knowledge required
  - Diversity indicates the scope and the breadth of knowledge domain
  - Distance gives the relationship of the concepts with the ontology root and a pseudo similarity measure

- Threshold coefficient plays important role
  - Coverage and diversity values change according to threshold coefficient
  - Threshold coefficient changes the projection graph to desired semantic significance

- Conceptual Distance behavior is same for changing threshold coefficient values as it is independent of the projection graph.
  - Gives an inverse similarity measure for subject concepts with respect to ontology root (rather than local root, for which definition can be easily extended).
Qualitative Data Analysis

- Questions are sorted according to those with high inverse correlation and those with lower inverse correlation between coverage-average score.
Questions sorted according to diversity
Correlation based analysis

- Large clustering (big circle)
- Dispersed concepts distribution and Diversity.
- Small Clustering
- Quiz based concepts distribution (200-400 and 750-1000)
- …more
Test based analysis

- most problems contain concepts in and around 200-400 and 700-1000
- concepts in problems go on increasing
- clustering denote projections of mapped concepts
Conclusions:

- For an automatic test design system and assessment framework is a must.

- To make course ware resources reusable and machine interpretable they have to represented in context. Semantic representation standards like RDF and OWL are used to represent this context.

- A representation language schema for course knowledge representation using ontology is given. The language is in OWL Lite and is expressible and computable.

- Problem complexity and knowledge content can be computed by applying synthetic parameters to course ontology having known the concept mapping. It is observed that the parameters are pretty good indicators of problem complexity.

- Assessment system can be intuitively be applied to automatic test design.
Related Work

- Problem assessment
  - Li, Sambasivam – Static knowledge structure
  - Rita Kuo et. al. – Information objects of simple questions

- Cognitive
  - Lee, Heyworth – Difficulty factors (perceived steps, students degree of familiarity, operations and expression in a problem)
  - Koedinger, Heffernan et.al. – number of symbols, ambiguous language
References:


Thank you.

Questions???