1. Introduction

**Observation:** Encyclopedic resources like Wikipedia have good reference value and broad coverage, but have limited pedagogic value. Textbooks on the other hand are often static and limited in coverage, but have limited pedagogic value. Textbooks and Wikipedia have good reference value and broad coverage.

2. Concept network

- Edge from a concept A to a concept B signifies that A is used to define B.
- Circularities in the concept network:
  - Description of Concept A assumes that Concept C is known.
  - Description of Concept C assumes that Concept A is known.
- Ideally, a pedagogic resource should ensure that the concept network is a directed acyclic graph.

3. Soft AND-OR composition

- User has to know either all $a_i$ s or all $b_j$ s to understand the Concept C. Soft AND-OR composition imposes relaxation on AND

4. Example

- Learning blanket encompasses the set of concepts in the concept graph that learner is familiar with.

5. Methods to resolve circular dependencies

- Perceptual grounding
  - Eg: Concepts like red can be defined by contrasting against other colors.

- Collapsing
  - Eg: Concepts such as polite and courteous can be defined using a single definition showing good manners.

- Linguistic grounding
  - Eg: The circular definition of opinion in the example depicted in section 5 can be broken by defining it as a personal view instead of the current definition a judgment of a person.

6. Greedy discovery of concepts for grounding

- Identify the set of concepts that do not take part in any cycle.
- Rank the remaining concepts based on the extent to which they affect learning due to cycles.
- Add concepts one by one in the ranked order, until there are no more cycles in the graph.

7. Ranking concepts for grounding

- **Relative Coverage**
  $$\text{Relative Coverage} \quad \text{helpsUnderstand}(a,b) \iff a \text{ occurs in the definition of } b$$
  $$\text{Coverage}(a) = \{|b| \text{ helpsUnderstand}(a,b)\}$$
  $$\text{Reachability}(b) = \{|a| \text{ helpsUnderstand}(a,b)\}$$
  $$\text{RelativeCoverage}(a) = \frac{\sum_b \text{Coverage}(a)}{\text{Reachability}(b)}$$

- **Weakness:** Ignores transitive closures and implicitly assumes OR composition.

- **PageRank**
  - Recursively estimate the importance of concepts in the concept network.
  - **Limitation:** Score of a concept increases (decreases) with increase (decrease) in the score of any of its in-neighbors.

- **Proposed Solution:** Weighted PageRank with weight as Relative Coverage.

8. Identifying regions for collapsing

- **Hypothesis:** Nodes in Strongly Connected Components (SCC) are related.
- **Collapse SCCs** in which the number of nodes is less than some threshhold.
- **Rank SCCs** using topological sort.

9. Resources

- **Learning resources:** content words in Brown and Gutenberg corpora.
  - Brown Corpus - 23,238 content words
  - Gutenberg Corpus - 18,361 content words

10. Comparison of methods

<table>
<thead>
<tr>
<th>Avg. level of learning blanket</th>
<th>Relative Coverage</th>
<th>PageRank</th>
<th>PageRank (Rel Cov)</th>
<th>Manual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>11.9</td>
<td>14.1</td>
<td>14.1</td>
<td>15.6</td>
</tr>
<tr>
<td>2</td>
<td>12.0</td>
<td>12.6</td>
<td>12.2</td>
<td>11.4</td>
</tr>
<tr>
<td>3</td>
<td>12.5</td>
<td>12.3</td>
<td>12.5</td>
<td>11.3</td>
</tr>
<tr>
<td>4</td>
<td>12.5</td>
<td>12.2</td>
<td>12.4</td>
<td>12.1</td>
</tr>
<tr>
<td>5</td>
<td>12.0</td>
<td>12.0</td>
<td>12.0</td>
<td>10.6</td>
</tr>
<tr>
<td>6</td>
<td>11.2</td>
<td>11.4</td>
<td>11.4</td>
<td>9.0</td>
</tr>
<tr>
<td>7</td>
<td>11.0</td>
<td>11.0</td>
<td>11.0</td>
<td>8.5</td>
</tr>
</tbody>
</table>

- Smaller fraction of concepts should be flagged for editing to reduce the human effort. Values in bold correspond to the best reductions.

11. Example - concepts suggested for collapsing

12. Conclusion and future work

- **Contribution:**
  - Identified characteristics of the encyclopedic resources that hinder learning.
  - Proposed approaches that help experts create a pedagogic view of an encyclopedic resource. Experiments show significant reduction in the number of concepts flagged for editing.

- **Future work:**
  - Explore personalized extensions to discover the learning blanket of a learner to help her explore the pedagogic space of concepts.
  - Extend to resources like Wikipedia.

**References**


**Contact:** ditty@cse.iitm.ac.in