Word Sense Disambiguation Using Semantic Graph
(Narayan Unny and Pushpak Bhattacharyya)

A presentation by
Ranjini Swaminathan
University of Arizona
Introduction

• One word – many senses
• Bank – financial institution, river, piggy bank
• DAG-Directed acyclic graph
• Method uses WordNet synsets
Text Structure

• *Constraint of proximity* – most synsets lie close together in the WordNet graph.

• *Conceptual density* – The number of words that lie in the sub hierarchy of synsets of the candidate word.

• *anaphors*
Algorithm

- Nodes → different senses of the words
- Arcs → dependency
- DAG constructed
- Start with monosems
Algorithm contd.

1. Collect monosems.
2. Initialize scores to 1.
3. Find link distance – breadth first search; *search-depth* is the cut-off.
4. Form the Semantic DAG –
   1. New node
   2. Child node (already present)-pointer added
Algorithm contd...

5. Pass the score of the parent to child

\[
\begin{align*}
\text{Score}(W_j^i) & \propto \text{Score}(W_k^{i-1}) \\
\text{Score}(W_j^i) & \propto \frac{1}{\text{Dist}(W_j^i, W_i^{i-1})} \\
\text{Score}(W_j^i) & \propto \frac{1}{\text{Link} \_ \text{dist}(W_j^i, W_i^{i-1})} \\
\text{Score}(W_j^i) & \propto \frac{\text{Score}(W_k^{i-1})}{\text{Dist}(W_j^i, W_k^{i-1}) \times \text{Link} \_ \text{dist}(W_j^i, W_k^{i-1})}
\end{align*}
\]

Score of parent
Distance in the text
Link distance in WordNet
Disambiguation

• Sense with the highest score taken.
• Every word (in a level) present only once in the DAG.
• Error cascading – cut-off level
Conclusions

• Performance results provided in terms of precision, recall and coverage.
• Performance varies with search depth and granularity of evaluation.
Further improvements and concerns

Improvements
• splitting text into segments for efficiency
• paragraph demarcation

Concerns
• error-cascading
• search–depth determination (abstract vs detailed text)