Nearest Neighbour Search with Zero-Suppressed Decision Diagram for Text

Yong Boon, Lim
Supervisor: Scott Sanner
Outline

- What’s Zero-Suppressed Decision Diagram (ZDD)
- Nearest Neighbour Search with ZDD.
- Goals
- Q & A
Zero-Suppressed Decision Diagram

- Compact way of representing sets of subsets.

- A variable is present if the path go through the 1-edge (solid line) of the node labelled with this variable.

- Example:
  \[ \{\{v3, v2, v1\}, \{v3, v2\}, \{v2, v1\}\} \]
Simple Example

- Each subset represents a document.
  - Doc1 = \{v3, v2, v1\}
  - Doc2 = \{v3, v2\}
  - Doc3 = \{v2, v1\}

- Which document is closest to Doc4 = \{v3, v1\}?

- Which document is closest to Doc5 = \{v2, v1\}?
Goals

- Implement Nearest Neighbour Search with Zero-Suppressed Decision Diagram.
- Use approximation techniques to speed up the search.
- Use complex similarity metrics beyond Hamming Distance.
- Compare with one other more “intelligent” NN search method.
Q & A