NGordNet (WordNet) CS 61B
Spring 2024

Ajay Sharma

Remark

If you would like to see the project source code, please contact me at ajay.sharma@berkeley.edu.

Overview

NgordNet is a graph-based application that analyzes word relationships through hierarchical structures and shortest-path computations. Using advanced data structures, the project enables efficient queries for synonyms, hypernyms, hyponyms, and other semantic connections. It demonstrates how graph algorithms, combined with effective data modeling, can process and navigate complex linguistic data efficiently and accurately. NgordNet bridges concepts in data structures and real-world natural language processing applications.

Description

NgordNet processes data sourced from WordNet, a comprehensive lexical database designed for natural language processing tasks. Words are represented as nodes, while edges encode semantic relationships such as synonyms, hypernyms, and hyponyms. The dataset is modeled as a directed acyclic graph (DAG) to capture hierarchical groupings and inheritance structures, ensuring flexibility in handling various types of linguistic queries.

Methodology

- **Data Structures**: Designed adjacency lists to efficiently represent graph edges and nodes, enabling scalability and rapid traversal for large datasets.
- **Algorithms**: Implemented Dijkstra's algorithm for shortest paths and traversal methods such as depth-first search (DFS) to navigate the graph.
- **Optimization**: Analyzed runtime complexity to improve performance, reducing query response times and ensuring smooth handling of large and complex graphs.
- **Error Handling**: Developed robust error-checking mechanisms for invalid queries, such as cyclic dependencies or disconnected nodes, ensuring stability and accuracy.

Results and Insights

NgordNet accurately retrieved semantic relationships, such as synonyms and hypernyms, by leveraging advanced graph algorithms. Its scalable design ensured efficient performance on large datasets without significant degradation. The project highlighted the effectiveness of combining data structure design with algorithmic problem-solving to process complex hierarchies, paving the way for further applications in linguistic analysis and graph-based systems.