Thesis Summary

Heather D. Pfeiffer

Title:

The Effect of Data Structures Modifications On Algorithms For Reasoning Operations

Problem Being Addressed:

Knowledge representation (KR) is used to store and retrieve informational data that is believed to contain knowledge. It is known that knowledge cannot be directly stored in the computer; therefore, a series of levels of representation is needed to transform the knowledge to a format that the computer can use. This transformed knowledge is saved in the machine using a dynamic data structure that is suitable for the style of KR being implemented. This data structure, together with the contents of the transformed knowledge, is known as the knowledge base. The KR allows the system to manipulate the knowledge in the data by using reasoning operations.

Importance of Problem:

Data structures and algorithms used in reasoning operations for operating over knowledge bases must be flexible, fast and efficient. The reasoning operation is the basis for querying the knowledge base, and as explained above uses a knowledge representation (KR). One such KR is "semantic networks" which represent a set of relationships among conceptual units in a discrete structure by mapping links between nodes in a net-like graph structure to render organized meaning (semantics). This dissertation addresses a particular style of semantic network KR known as Conceptual Graphs (CGs). Conceptual graphs are connected, directed, bipartite graphs, where the bipartite objects are divided into concepts and conceptual relationships. A user application or system can store and retrieve from the knowledge base (KB) by using query graphs through a generic interface. The queries are given the same data structure as the knowledge base and changing the data structure of both the knowledge base and queries can allow the reasoning operation to be performed more quickly and execute more efficiently.

Research Plan:

An algorithm for performing the reasoning operation is dependent on the data structure it uses. In this dissertation, the two basic reasoning algorithms, *maximum projection* and *maximum join*, will be defined. I will show that these algorithms (for best case) execute in polynomial time when using a CG knowledge representation. The maximum projection operation is the matching of a labeled subgraph isomorphicly onto another labeled graph; however, these graphs contain type information that can be generalized and specialized using a lattice ordering in the related 'type hierarchy'. Maximum join is a basic graph matching algorithm again using the type information during the matching of nodes.

After defining the projection algorithm, standard basic data structures are analyzed theoretically and experimentally modified. It will be shown that altering the internal data structure of the knowledge base improves the running time of the maximum projection operation from n^3 for the defined algorithm to n^2 in the best case and $O(n^2 \log n)$ for the average case where n is the maximum number of nodes in a graph. The runtime time efficiency of the related reasoning operation, maximum join, is examined after the evaluation of the maximum projection operation is completed.

A generic interface will be made available to the application experimentally implementing the reasoning operations. This interface will be tested using an outside user application.

Progress Made and To Be Made:

Projection algorithm analysis

Theoretical analysis of the maximum projection algorithm has been done and is being compared to three other known projection algorithms. This analysis will be presented during the thesis proposal presentation. The experimental modification of the data structures has produced preliminary results; more extensive experimental testing will be performed on the data structures for the projection operation.

Maximum join analysis

The maximum join algorithm is defined and theoretical analysis is proceeding. Research is being performed for comparable known algorithms to do theoretical analysis against. When the projection algorithm is fully analyzed and tested, the same analysis and testing will be performed on the maximum join algorithm.

Generic interface

The generic interface is designed and implemented and is currently being tested with a user application. Again this data will be presented at the thesis proposal presentation.

Work

Done:

Define projection algorithm; theoretical analysis of projection operation; preliminary results of projection operation; define maximum join algorithm; generic interface to experimental system.

Remains:

Complete experimental analysis of maximum projection operation; theoretical analysis of maximum join operation; experimental analysis of maximum join operation; and improvements to the generic interface to experimental system.

Thesis Sentence:

This dissertation addresses both the efficiency of the underlying algorithms for knowledge reasoning given a particular knowledge representation where the data structures and algorithms for the operations of projection and maximal join are examined, and provides a modular flexible generic interface for using these efficient reasoning operations for storage and retrieval to a knowledge base.