Wang, P. (2000). Comparing cognitive maps using graph algorithms. . 11th ASIS SIG/CR Classification Research Workshop, 157-159. doi:10.7152/acro.v11i1.12781

Comparing Cognitive Maps Using Graph Algorithms

A topic for the Idea Mart
Classification for User Support and Learning
11th ASIS SIGCR Workshop

Peiling Wang
University of Tennessee
peilingw@utk.edu

INTRODUCTION

The purpose of this project is two-fold: (1) to examine researchers' knowledge structures on research topics; (2) to compare knowledge structures of experts with those of non-experts. Expert is defined as the researcher who had conducted in-depth research and published on the topic and whose vocabulary used to describe the topic for online searching was the basis for constructing maps. Non-experts are also researchers in the same field but had not done any in-depth work on the topics when they were asked to make a map using the given vocabulary. Both experts and non-experts were allowed to add new terms to or drop original terms from the given vocabulary. The finished cognitive map is a structured layout of the terms on a two-dimensional plane. During the mapping process, subjects also provided thinking-aloud protocols, which revealed additional information on how they saw the relationships of the concepts represented by the terms.

Preliminary analyses of ten sets of cognitive maps for ten research topics revealed differences in final vocabulary (after adding and dropping terms), configuration (top-down, left-right, radial, etc.), and foci (focusing on problems, issues or processes). These results were reported at the 1999 ASIS Annual Meeting.

Work has just been completed to advance the comparison of semantic closeness using graph theory to convert maps into matrices and to calculate similarities. The following algorithms have been developed for the conversion and calculation of cognitive maps.

ALGORITHAMS

1. Adjacency matrix A

This matrix is used to represent nearest neighbors of a term:
a_i represents a term on the map (in alphabetical order)
a_{ij} is used to store adjacent terms in a matrix
$a_{ij} = 1$ if directly linked and $a_{ij} = 0$ otherwise
$0 =< i <= N-1$; $0 =< j <= N-1$
N: number of terms in the vocabulary; the union of terms in the set of maps
Only lower triangular matrix is stored.
Note: synonyms form cliques in a graph

ISSN: 2324-9773

Wang, P. (2000). Comparing cognitive maps using graph algorithms. . 11th ASIS SIG/CR Classification Research Workshop, 157-159. doi:10.7152/acro.v11i1.12781

2. Distance matrix D

This matrix establishes the path among the related terms. An algorithm is used to scan matrix A to calculate the shortest distance between each pair of related terms. The matrix D records the shortest path (minimum links/hops) between each pair of terms:

$d_{ij} = n$
n is the shortest path between $term_i$ and $term_j$
$0 =< n <= N-1$ (n = 0 if the two terms are not related)
N: number of terms in the vocabulary; the union of terms in the set of maps

3. Submatrices

Concepts are clustered on cognitive maps. An algorithm has been implemented to find unique clusters in matrix D. A cluster is a connected submatrix representing a subgraph of the graph.

4. Similarity of clusters

Similarity of clusters from each set of maps is calculated using the following algorithm:

Similarity ($Cluster_{map1}$, $Cluster_{map2}$) = C / (A+B-C)

C: number of common terms
A, B: number of terms in $Cluster_{map1}$ and $Cluster_{map2}$ respectively

0<= Similarity <= 1 (0: completely different; 1: exactly the same)

This is a modification of Dice coefficient used for document clustering (Frakes and Baeza-Yates, 1992, p. 422):

Similarity (document1, document2) = 2C / (A+B)

C: number of terms in common
A, B: number of terms in document 1, 2 respectively

5. Semantic differences of clusters with high similarity values

Structural differences of common terms (C from 4) are compared to identify differences between maps. For example, within a set of common terms, terms may be related in different ways reflecting depth or breadth of one's knowledge.

CONCLUSION

In summary, the algorithms based on graph theory developed in this project have successfully converted cognitive maps to matrices and calculated clusters, similarities, and differences; the results are very exciting. A cognitive map can be represented by a graph (nodes for concepts and links for associations), which is conveniently represented by a matrix (two dimensional data structure). Graph-based representation of knowledge structure has a great potential in information retrieval, which is worth further exploration and experimentation.

[July 20, 2000]

ISSN: 2324-9773

Wang, P. (2000). Comparing cognitive maps using graph algorithms. . 11th ASIS SIG/CR Classification Research Workshop, 157-159. doi:10.7152/acro.v11i1.12781

Questions

1. Do the cognitive maps drawn by the subjects represent natural organizations of concepts? How can we find answers to this question?

2. Is the graph-based approach to representation of information needs an effective method as compared to Boolean-based queries?

3. It may not be practical to represent a large document collection using graph-based approach. How can graph-based queries be used in a collection that is indexed using traditional method?

ISSN: 2324-9773