Comparision Between Features of CbO based Algorithms for Generating Formal Concepts

Nuwan Kodagoda, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka

Koliya Pulasinghe, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka

ABSTRACT

Formal Concept Analysis provides the mathematical notations for representing concepts and concept hierarchies making use of order and lattice theory. This has now been used in numerous applications which include software engineering, linguistics, sociology, information sciences, information technology, genetics, biology and in engineering. The algorithms derived from Kustenskov's CbO were found to provide the most efficient means of computing formal concepts in several research papers. In this paper key enhancements to the original CbO algorithms are discussed in detail. The effects of these key features are presented in both isolation and combination. Eight different variations of the CbO algorithms highlighting the key features were compared in a level playing field by presenting them using the same notation and implementing them from the notation in the same way. The three main enhancements considered are the partial closure with incremental closure of intents, inherited canonicity test failures and using a combined depth first and breadth first search. The algorithms were implemented in an un-optimized way to focus on the comparison on the algorithms themselves and not on any efficiencies provided by optimizing code. The main contribution of this paper is the complete comparison of the three main enhancements used in recent variations of the CbO based algorithms. The main findings were that there is a significant performance improvement partial closure with incremental closure of intents is used in isolation. However, there is no significant performance improvement when the depth and breadth first search or the inherited canonicity test failure feature is used in isolation. The inherited canonicity test failure needs to be combined with the combined depth and breadth first feature to obtain a performance increase. Combining all the three enhancements brought the best performance.

KEYWORDS

CbO, FcbO, Formal Concept Analysis, In-Close, Knowledge

BACKGROUND

Knowledge is represented in Formal Concept Analysis as a formal context. This describes a binary relationship between a set of objects and a set of attributes of a domain.

A precise definition of a formal context is given below.
A formal context is defined as $K := \left(G, \ M, I \right)$

Where G is a set of objects, M a set of attributes and I a binary incidence relationship between G and M with $I \subseteq G \times M$

Since formal contexts are a binary relationship they can be represented as cross tables. Here each object and attribute is represented as a row and a column respectively (Wille, 1982; see also Ganter, Stumme, & Wille, 2002)

Let's take an example to illustrate formal concept definitions making use of the Solar system. Figure 1 shows the eight major planets in our Solar System with details of seven different properties. The table has the eight planet as rows and the eight properties as columns. A cross is used to mark the presence of a property against a planet. For example, the planets Jupiter, Saturn, Uranus and Neptune are considered to be gas giant planets. This is denoted in the table with crosses marked against each of the above planets for the property gas giant planet. The rows are known as extents and the columns are known as intents. A formal concept is defined as a pair of maximum of a set of extents and a set of intents. For the earlier example, we can also see that the four planets also have the property Has Some Moon checked. So the formal concept consisting of gas giants is defined as ({Jupiter, Saturn, Uranus, Neptune}, {Gas Giant, Has Some Moon}). Similarly, we can define a formal concept of planets that are gas giants and where the sun sets in the West. From Figure 1, we can find the following concept ({Jupiter, Saturn, Neptune}, {Gas Giant, Sun sets in the west, Has some Moon}). Here too the property Has some moon is a common property and is included as part of the formal concept. The planets Jupiter, Saturn and Neptune are the gas giant planets where the sun sets in the west. The planets mentioned above are the extent of this formal concept and the three attributes form the intent. A mathematical definition of Formal Concepts is given below.

For a set of objects $A \subseteq G$ the set A' is defined as

$$A' := \left\{ m \in M \mid m\ I\ g \ \text{ for all } g \in A \right\}$$

Similarly, for a set of attributes $B \subseteq G$ the set B' is defined as

$$B' := \left\{ g \in G \mid m\ I\ g \ \text{ for all } m \in B \right\}$$

(A, B) is a formal concept if $A' = B$ and $B' = A$.

There are many formal concepts in a formal context. All the possible formal concepts that are there in a formal context can be generated and be represented in a concept hierarchy. A concept lattice can be used to represent a concept hierarchy (See Figure 2). Here the extent and intent indices *1* to *8* and *a* to *g* are used for brevity. The two examples mentioned earlier map to the concepts *({5,6,7,8},{a,b})* and *({5,6,8},{a,b,c})* respectively. They are the left most vertex of the lattice and the vertex immediately below. In the example taken the concept *({5,6,7,8},{a,b})* is the sub concept of *({5,6,8},{a,b,c})*. Accordingly in the concept lattice the concept *({5,6,7,8},{a,b})* is above the concept *({5,6,8},{a,b,c})* as it is more general. A mathematical definition of the relationships between formal concepts is given below.

Let $K := (G, M, I)$ define a formal context. Assuming (A_1, B_1) and (A_2, B_2) are formal concepts of K.

$$(A_1, B_1) \leq (A_2, B_2) := A_1 \subseteq A_2 \left(\Leftrightarrow B_1 \supseteq B_2 \right)$$

(A_1, B_1) is called the sub concept of (A_2, B_2) and (A_2, B_2) is called the super concept of (A_1, B_1) In the Concept Lattice the more general concepts are above and the specific concepts are below.

The set of all formal concepts of K together with the defined order relation is denoted by $\beta(G, M, I)$ which can be represented as a Concept Lattice (See Figure 2).

Related Content

Fuzzy Lattice Ordered G-modules
Ursala Pauland Paul Isaac (2019). *International Journal of Fuzzy System Applications (pp. 94-107).*
www.irma-international.org/article/fuzzy-lattice-ordered-g-modules/233588

Logarithmic Entropy Measures for Fuzzy Rough Set and their Application in Decision Making Problem
Omdutt Sharmaand Priti Gupta (2020). *International Journal of Fuzzy System Applications (pp. 80-97).*
www.irma-international.org/article/logarithmic-entropy-measures-for-fuzzy-rough-set-and-their-application-in-decision-making-problem/250821

Inference Strategies for Solving Semi-Markov Decision Processes
Matthew Hoffmanand Nando de Freitas (2012). *Decision Theory Models for Applications in Artificial Intelligence: Concepts and Solutions (pp. 82-96).*
www.irma-international.org/chapter/inference-strategies-solving-semi-markov/60925

A Combined Fuzzy Method for Evaluating Criteria in Enterprise Resource Planning Implementation
Hodjat Hamidi (2016). *International Journal of Intelligent Information Technologies (pp. 25-52).*
www.irma-international.org/article/a-combined-fuzzy-method-for-evaluating-criteria-in-enterprise-resource-planning-implementation/152304

ISCG: An Intelligent Sensing and Caption Generation System for Object Detection and Captioning Using Deep Learning
Aahan Singh, Nithin Nagaraj, Srinidhi Hiriyannaiahand Lalit Mohan Patnaik (2020). *International Journal of Intelligent Information Technologies (pp. 51-67).*
www.irma-international.org/article/iscg/262979