

Deriving Event Relevance from the Ontology Constructed with Formal Concept Analysis

Wei Xu^{1,2}, Wenjie Li¹, Mingli Wu¹, Wei Li¹, Chunfa Yuan²

¹ Department of Computing,
The Hong Kong Polytechnic University, Hong Kong
{cswxu, cswli, csmlwu, cswli}@comp.polyu.edu.hk

² Department of Computer Science and Technology
Tsinghua University, China
{vivian00, cfyuan}@mails.tsinghua.edu.cn

Abstract. In this paper, we present a novel approach to derive event relevance from event ontology constructed with Formal Concept Analysis (FCA), a mathematical approach to data analysis and knowledge representation. The ontology is built from a set of relevant documents and according to the named entities associated to the events. Various relevance measures are explored, from binary to scaled, and from symmetrical to asymmetrical associations. We then apply the derived event relevance to the task of multi-document summarization. The experiments on DUC 2004 data set show that the relevant-event-based approaches outperform the independent-event-based approach.

1 Introduction

Extractive summarization is to select the sentences which contain salient concepts in documents. An important issue with it is what criteria should be used to extract the sentences. Event-based summarization attempts to select and organize the sentences in a summary with respect to the events or the sub-events that the sentences describe [1, 2]. As the relevance of events reveals the significance of events, it helps singling out the sentences with the most core events. However, the event-based summarization techniques reported so far explored the events independently.

In the realm of information retrieval, term relations were commonly derived either from a thesaurus like WordNet or from the corpus where the contexts of the terms were investigated. Likewise, mining event relevance requires taking contexts of event happenings into account. The event contexts in our definition are event arguments, such as participants, locations and occurrence times, etc. They are important in defining events and distinguishing them from one another. By this observation, we make use of the named entities associated with the events as event contexts and characterize the events with the verbs and action-denoting nouns prescribed by the named entities.