Chapter 19 The Role of Textual Graph Patterns in Discovering Event Causality

Bryan Rink University of Texas at Dallas, USA

Cosmin Adrian Bejan University of Southern California, USA

Sanda Harabagiu University of Texas at Dallas, USA

ABSTRACT

We present a novel method for discovering causal relations between events encoded in text. In order to determine if two events from the same sentence are in a causal relation or not, we first build a graph representation of the sentence that encodes lexical, syntactic, and semantic information. From such graph representations we automatically extract multiple graph patterns (or subgraphs). The patterns are sorted according to their contribution to the expression of intra-sentential causality between events. To decide whether a pair of events is in a causal relation, we employ a binary classifier that uses the graph patterns. Our experimental results indicate that capturing causal event relations using graph patterns outperforms existing methods.

INTRODUCTION

Automatic discovery of causal relations between textual events is a central task for various applications in Natural Language Processing (NLP). Specifically, tasks that require some form of reasoning such as probabilistic reasoning (Narayanan, 1997), common sense reasoning (Mueller,

DOI: 10.4018/978-1-60960-741-8.ch019

2007), and question answering (Girju, 2003). An example of a causal relation between two events is presented in the following sentence:

• **S1:***Hewoundup* and*letloose*a fastball.

This example encodes a special type of causal relation called ENABLEMENT, in which the event *wound* enables the event *let loose* to happen. In this case, the event *let loose* will take place only

if the event *wound* occurs. Moreover, causal relations are closely related with temporal relations that hold between two events. For example, the ENABLEMENT relation from sentence S1 corresponds to a BEFORE temporal relation established between the *wound* event and the *let loose* event.

We cast causal relation detection as a classification task. Most existing approaches for discovering causal relations use machine learning classifiers based on unstructured linguistic features. In this chapter, we describe a method that captures the contextual information of the two events. Contextual information is captured by lexical, syntactic, and semantic features as well as dependencies that exist between these features. Features are structured into graphs, enabling us to discover graph patterns that account for causality between pairs of intra-sentential events.

The remainder of the chapter is organized as follows. We first describe the related work. Next we detail the graph representation that captures the structural dependencies of lexical, syntactic, and semantic features characterizing pairs of events that belong to a causal relation. In that section we detail each type of feature as well. The following section discusses the method of discovering syntactic and lexico-semantic graph patterns that characterize causality. Patterns are used for training a binary classifier which decides whether there exists a causal relation between pairs of intra-sentential events. In a separate section we discuss the experimental results and the error analysis. We also discuss the future research directions before summarizing the conclusions of the chapter.

RELATED WORK

Causality can be expressed in texts in many ways. Moreover, linguists and researchers have established that there are several forms of causality, including PURPOSE, REASON, and ENABLEMENT (Bejan & Harabagiu, 2008a). Traditionally, causality in text has been viewed as a discourse phenomenon or as a semantic-specific function. Methods that have focused on the discourse-level information attempted to recognized causal relations between discourse units in the same way as they would recognize other discourse relations (e.g. ELABO-RATION, PARALLEL, and VIOLATED-EXPECTATIONS). For example, Marcu & Echihabi (2002) have used a naïve Bayes classifier trained on pairs of words from discourse segments separated by cue phrases such as *thus* and *because*. Their approach achieved very high scores when evaluated on CAUSE-EXPLANATION-EVIDENCE relations.

The previous techniques that focused on the semantic aspects of causality have relied on a combination of syntactic and lexico-semantic knowledge. For example, Girju and Moldovan (2002) have studied causal relations between nominals. They were able to obtain promising results by focusing on patterns of the form [NOUN-VERB-NOUN] and using semantic knowledge encoded in WordNet (Fellbaum et al., 1998). WordNet is a large lexico-semantic knowledge base that encodes the vast majority of English nouns, verbs, adjectives, and adverbs. Word senses and hypernym information available from WordNet were used to detect causal relations in the approach reported in (Girju and Moldovan 2002). The results were promising, and thus they have prompted a task dedicated to the recognition of nominal causal relations in SemEval 2007. SemEval (formerly Senseval) is a series of evaluations of semantic analysis systems. The task for classification of semantic relations between nominals asked participants to identify whether a given semantic relation (e.g., PRODUCT-PRODUCER, PART-WHOLE, CAUSE-EFFECT, etc.) exists between two nouns from a sentence. The approach reported in (Beamer et al., 2007) uses several lexico-syntactic features, including the word stems and sequences of part of speech between the nominals. They used some general features including whether the nouns encode a time or location. They also used features based on the context of the nouns such as gram15 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/role-textual-graph-patterns-discovering/61057

Related Content

Adaptive Filters: Structures, Algorithms, and Applications

Sergio L. Nettoand Luiz W.P. Biscainho (2007). Advances in Audio and Speech Signal Processing: Technologies and Applications (pp. 190-224). www.irma-international.org/chapter/adaptive-filters-structures-algorithms-applications/4687

Semantics of Techno-Social Spaces

Sergey Maruev, Dmitry Stefanovskyiand Alexander Troussov (2015). *Modern Computational Models of Semantic Discovery in Natural Language (pp. 204-234).* www.irma-international.org/chapter/semantics-of-techno-social-spaces/133880

DSP Techniques for Sound Enhancement of Old Recordings

Paulo A.A. Esquefand Luiz W.P. Biscainho (2007). *Advances in Audio and Speech Signal Processing: Technologies and Applications (pp. 93-130).* www.irma-international.org/chapter/dsp-techniques-sound-enhancement-old/4684

Maximizing ANLP Evaluation: Harmonizing Flawed Input

Adam Renner, Philip M. McCarthy, Chutima Boonthum-Deneckeand Danielle S. McNamara (2012). *Applied Natural Language Processing: Identification, Investigation and Resolution (pp. 438-456).* www.irma-international.org/chapter/maximizing-anlp-evaluation/61064

Co-Designing Participatory Tools for a New Age: A Proposal for Combining Collective and Artificial Intelligences

José Luis Fernández-Martínez, Maite López-Sánchez, Juan Antonio Rodríguez Aguilar, Dionisio Sánchez Rubioand Berenice Zambrano Nemegyei (2020). *Natural Language Processing: Concepts, Methodologies, Tools, and Applications (pp. 860-877).*

www.irma-international.org/chapter/co-designing-participatory-tools-for-a-new-age/239970