Semantic Understanding of Professional Soccer Commentaries

Hannaneh Hajishirzi, Mohammad Rastegari, Ali Farhadi, and Jessica K. Hodgins
{hannaneh.hajishirzi, mrastega, afarhadi, jkh}@cs.cmu.edu
Disney Research Pittsburgh and Carnegie Mellon University

Abstract

This paper presents a novel approach to the problem of semantic parsing via learning the correspondences between complex sentences and rich sets of events. Our main intuition is that correct correspondences tend to occur more frequently. Our model benefits from a discriminative notion of similarity to learn the correspondence between sentence and an event and a ranking machinery that scores the popularity of each correspondence. Our method can discover a group of events (called macro-events) that best describes a sentence. We evaluate our method on our novel dataset of professional soccer commentaries. The empirical results show that our method significantly outperforms the state-of-the-art.

1 Introduction

This paper addresses the problem of understanding professional commentaries of soccer games. Computational understanding of such domains has crucial impact in automatic generation of commentaries and also in game analysis and strategic planning. To this end, one needs to infer the semantics of natural language text; this is an extremely challenging problem. Understanding professional soccer commentaries further introduces interesting and challenging issues. For example, commentators do not typically talk about all the events of the game, selecting what is important. Also, they use a variety of phrases to report similar events. For example, a simple event of “A passes to B” can be commentated in several different ways: “A feeds B”, “A and B in a nice combination”, “A, what a beautiful way to B”. Further, in some cases, commentators create a group of events and only mention a macro-event. For example, instead of saying “A passes to B, B passes to C, and C passes to D”, the commentators report this whole sequence of events as “Team X is coming forward” or “nice attack by X”. Also, professional commentators report several statistics and related information about the league, players, stadium, and weather during less interesting segments of a play.

A general solution to understanding such a complex phenomenon requires inferring about game-related events, reasoning in terms of very complex paraphrases, and also forming high-level understandings of game events. Most recent work in semantic parsing of natural language translates individual sentences into the underlying meaning representations. Meaning representations are usually logical forms represented with events or relations among entities. The problem of semantic parsing can be formulated as learning to map between sentences and meaning representation in a supervised fashion [Zettlemoyer and Collins, 2005]. One can decrease the amount of supervision in specific controlled domains, such as RoboCup soccer [Chen et al., 2010; Hajishirzi et al., 2011] and Windows help instructions [Branavan et al., 2009]. Recently, [Liang et al., 2009] introduce a general semantic parsing technique that is not restricted to a specific domain, but is not scalable to large datasets due to the complexity of the model. In this paper, we introduce an algorithm that does not require domain-specific knowledge and is scalable to larger datasets.

We formulate the problem of understanding soccer commentaries as learning to align sentences in commentaries to a list of events in the corresponding soccer game. Our approach does not need expensive supervision in terms of correspondences between sentences and events. Similar to previous work [Liang et al., 2009; Chen et al., 2010], we use loose temporal alignments between sentences in commentaries and events of games. We pair sentences with several events that occur in the rough temporal vicinities of the sentences. Each pair consists of a sentence and a corresponding event. We then try to distinguish between correct and incorrect pairs. We rank pairs based on how consistently they appear in other places. We use a discriminative notion of similarity to reason about repetitions of pairs of sentences and events. The core intuition is that,