Available online at www.sciencedirect.com

ScienceDirect

Computer Speech and Language 29 (2015) 1–19

COMPUTER SPEECH AND LANGUAGE

www.elsevier.com/locate/csl

Leveraging social Q&A collections for improving complex question answering☆

Youzheng Wu*, Chiori Hori, Hideki Kashioka, Hisashi Kawai

Spoken Language Communication Laboratory, National Institute of Information and Communications Technology, 3-5 Hikari-dai, Seika-cho, Soraku-gun, Kyoto 619-0289, Japan

Received 24 May 2012; received in revised form 23 May 2014; accepted 4 June 2014
Available online 12 June 2014

Abstract

This paper regards social question-and-answer (Q&A) collections such as Yahoo! Answers as knowledge repositories and investigates techniques to mine knowledge from them to improve sentence-based complex question answering (QA) systems. Specifically, we present a question-type-specific method (QTSM) that extracts question-type-dependent cue expressions from social Q&A pairs in which the question types are the same as the submitted questions. We compare our approach with the question-specific and monolingual translation-based methods presented in previous works. The question-specific method (QSM) extracts question-dependent answer words from social Q&A pairs in which the questions resemble the submitted question. The monolingual translation-based method (MTM) learns word-to-word translation probabilities from all of the social Q&A pairs without considering the question or its type. Experiments on the extension of the NTCIR 2008 Chinese test data set demonstrate that our models that exploit social Q&A collections are significantly more effective than baseline methods such as LexRank. The performance ranking of these methods is QTSM > {QSM, MTM}. The largest F_3 improvements in our proposed QTSM over QSM and MTM reach 6.0% and 5.8%, respectively.

Keywords: Complex question answering; Web mining; Summarization

1. Introduction

Unlike conventional search engines, which find relevant documents on the web, question-answering (QA) systems are designed to return much more focused answers, for example:

Q: In which city will the 2020 Olympic games be held?
A: Tokyo

☆ This paper has been recommended for acceptance by E. Briscoe.

* Corresponding author. Tel.: +81 774 98 6333; fax: +81 774 98 6823.
E-mail addresses: youzheng.wu@nict.go.jp, erzhengcn@gmail.com (Y. Wu), chiori.hori@nict.go.jp (C. Hori), hideki.kashioka@nict.go.jp (H. Kashioka), hisashi.kawai@nict.go.jp (H. Kawai).

http://dx.doi.org/10.1016/j.csl.2014.06.001

QA research attempts to deal with a wide range of question types, including factoid, list,[1] and complex questions. In the past, substantial progress has been made with factoid and list questions. In February 2011, IBM's Watson QA system (Ferrucci et al., 2010) defeated human grand champions in the game of Jeopardy![2] In this paper, we focus on complex questions because – apart from definitional[3] (Voorhees, 2003; Mitamura et al., 2008), reason (Higashinaka and Isozaki, 2008; Oh et al., 2013), and opinion (Dang, 2008) complex questions – many types remain largely unexplored.[4]

Compared to factoid and list questions that can be answered by short phrases, such as persons, organizations, locations and dates, complex questions, whose answers generally consist of a list of "nuggets" (Voorhees, 2003; Mitamura et al., 2008), are more difficult to answer. For instance, the factoid question, "In which city will the 2020 Olympic games be held?" asks for a city name, and thus it is easier to impose some constraints on the plausible answers for this type of question and significantly reduce the search space of plausible answers. However, complex questions such as "What are the hazards of global warming?" often seek multiple, different types of information simultaneously, making it difficult to screen plausible answers. Moreover, complex QA tasks require inferring and synthesizing information from multiple documents to provide multiple nuggets as answers (Dang, 2006; Chali et al., 2009). To answer complex questions, we often need to go through complex procedures.

Many approaches have been proposed to answer factoid, definitional, reason, and opinion questions. Among them, machine learning techniques have proven to be effective in constructing QA components from scratch, but these supervised techniques require a certain quantity of question and answer (Q&A) pairs as training data. For example, Echihabi and Marcu (2003) and Sasaki (2005) constructed 90,000 English and 2000 Japanese Q&A pairs for their factoid QA systems, respectively. Cui et al. (2004) collected 76 term-definition pairs for their definitional QA system. Higashinaka and Isozaki (2008) used 4849 positive and 521,177 negative examples in their reason QA system. Stoyanov et al. (2005) required a known subjective vocabulary for their opinion QA system. To answer other types of complex questions using supervised techniques, we need to collect Q&A pairs for each type of complex question to train models, even though this is an extremely expensive and labor-intensive task. Fortunately, many user-generated Q&A pairs can be found in social QA websites such as Yahoo! Answers,[5] Baidu Zhidao,[6] and Answers.com.[7]

This paper explores the automatic learning of Q&A training pairs and the mining of needed knowledge from social Q&A collections such as Yahoo! Answers. We are interested in whether millions of typically noisy user-generated Q&A pairs can be exploited for automatic QA systems. If so, a plethora of Q&A training data is already readily available.

Many studies (Riezler et al., 2007; Surdeanu et al., 2008; Duan et al., 2008; Wang et al., 2010) have retrieved similar Q&A pairs from social QA websites as answers to test questions; accordingly, answers cannot be generated for questions that have not previously been answered on such sites. Our study, however, regards social Q&A websites as knowledge repositories and exploits their knowledge to synthesize answers to questions that have not yet been answered. Even for questions that have been answered, it is necessary to perform answer summarization (Liu et al., 2008; Chan et al., 2012). Our approach can also be used for this purpose. To the best of our knowledge, very few works in the literature have addressed this aspect.

Various kinds of knowledge can be mined from social Q&A collections to support complex QA systems. In this paper, we present a question-type-specific method (QTSM) to mine question-type-specific knowledge and compare it with the question-specific and monolingual translation-based methods proposed in related work. Given question Q, whose question type Q_t is automatically recognized from Q, three kinds of methods can be applied here.

- Our proposed QTSM collects Q&A pairs in which the question types are the same as Q_t and extracts salient cue expressions that are indicative of possible answers to question type Q_t. It uses the expressions and the Q&A pairs to train a binary classifier to remove noisy candidate answers.

[1] Example question: Name 20 countries other than the United States that have a McDonald's restaurant.
[2] http://www.jeopardy.com/.
[3] Example questions: Who is Aaron Copland? What is a golden parachute?
[4] Most complex questions have generally been referred to as what-questions in previous studies. This paper argues that it is helpful to treat them discriminatively.
[5] http://answers.yahoo.com/.
[6] http://zhidao.baidu.com/.
[7] http://www.answers.com/.