



Multiple topic identification in human/human conversations[☆]

X. Bost^a, G. Senay^{a,*}, M. El-Bèze^a, R. De Mori^{a,b}

^a *LIA, University of Avignon, 339 chemin des Meinajariès, 84000 Avignon, France*

^b *McGill University, School of Computer Science, Montréal, Québec, Canada*

Received 9 January 2014; received in revised form 19 December 2014; accepted 13 March 2015

Available online 8 April 2015

Abstract

The paper deals with the automatic analysis of real-life telephone conversations between an agent and a customer of a customer care service (CCS). The application domain is the public transportation system in Paris and the purpose is to collect statistics about customer problems in order to monitor the service and decide priorities on the intervention for improving user satisfaction.

Of primary importance for the analysis is the detection of themes that are the object of customer problems. Themes are defined in the application requirements and are part of the application ontology that is implicit in the CCS documentation.

Due to variety of customer population, the structure of conversations with an agent is unpredictable. A conversation may be about one or more themes. Theme mentions can be interleaved with mentions of facts that are irrelevant for the application purpose. Furthermore, in certain conversations theme mentions are localized in specific conversation segments while in other conversations mentions cannot be localized. As a consequence, approaches to feature extraction with and without mention localization are considered.

Application domain relevant themes identified by an automatic procedure are expressed by specific sentences whose words are hypothesized by an automatic speech recognition (ASR) system. The ASR system is error prone. The word error rates can be very high for many reasons. Among them it is worth mentioning unpredictable background noise, speaker accent, and various types of speech disfluencies.

As the application task requires the composition of proportions of theme mentions, a sequential decision strategy is introduced in this paper for performing a survey of the large amount of conversations made available in a given time period. The strategy has to sample the conversations to form a survey containing enough data analyzed with high accuracy so that proportions can be estimated with sufficient accuracy.

Due to the unpredictable type of theme mentions, it is appropriate to consider methods for theme hypothesization based on global as well as local feature extraction. Two systems based on each type of feature extraction will be considered by the strategy. One of the four methods is novel. It is based on a new definition of density of theme mentions and on the localization of high density zones whose boundaries do not need to be precisely detected.

The sequential decision strategy starts by grouping theme hypotheses into sets of different expected accuracy and coverage levels. For those sets for which accuracy can be improved with a consequent increase of coverage a new system with new features is introduced. Its execution is triggered only when specific preconditions are met on the hypotheses generated by the basic four systems.

[☆] This paper has been recommended for acceptance by T. Kawahara.

* Corresponding author. Tel.: +33 14083879542.

E-mail addresses: xavier.bost@univ-avignon.fr (X. Bost), gregory.senay@gmail.com (G. Senay), marc.elbeze@univ-avignon.fr (M. El-Bèze), rdemori@cs.mcgill.ca (R. De Mori).

Experimental results are provided on a corpus collected in the call center of the Paris transportation system known as RATP. The results show that surveys with high accuracy and coverage can be composed with the proposed strategy and systems. This makes it possible to apply a previously published proportion estimation approach that takes into account hypothesization errors.
© 2015 Elsevier Ltd. All rights reserved.

Keywords: Human/human conversation analysis; Multi-topic identification; Spoken language understanding; Interpretation strategies

1. Introduction

In recent years, there has been an increasing research interest in the analysis of human/human spoken conversations. Advances in this research area are described in [Tur and De Mori \(2011\)](#) and [Tur and Hakkani-Tür \(2011\)](#). A scientifically interesting and practically important component of this research is topic identification for which an ample review of the state of the art can be found in [Hazen \(2011\)](#). In spite of the relevant progress achieved so far, it is difficult to reliably identify multiple topics in real-life telephone conversations between casual speakers in unpredictable acoustic environments. Of particular interest are call-center conversations in which customers discuss problems in specific domains with an advisor. This is the case of the application considered in this paper. The purpose of the application is to monitor the effectiveness of a customer care service (CCS) of the RATP Paris transportation system by analyzing real-world human/human telephone conversations in which an agent solicits a customer to formulate a problem and attempts to solve it. The application task is to perform surveys of customer problems in different time periods. Proportions of problem themes computed with the survey data are used for monitoring user satisfaction and establishing priorities of problem solving interventions. Of primary importance for this task is the ability to automatically select, in a given period of time, a sufficiently large sample of conversations automatically analyzed with high accuracy. Application relevant information for evaluating proportions of problem items and solutions is described in the application requirements making evident, not formally defined, but useful, speech analytics. Themes mentioned in the application documentation appear to be the most important and general classes representing concerns that the survey has to address. Themes appear to be an adequate semantic representation of the types of problems in the application ontology that is inferred from the documentation.

Agents follow a pre-defined protocol to propose solutions to user problems about the transportation system and its services. In order to evaluate proportions of problem types, agents compile conversation reports following a domain ontology. Due to time constraints, reports compiled by the agents cover a small proportion of conversations. Furthermore, compiled reports are often incomplete and error-prone. Thus, an automatic classification of conversation themes would be useful for producing a quantity of accurate reports sufficient for performing a reliable survey.

The achievement of acceptable automatic solutions has to overcome problems and limits of known spoken language processing and interpretation systems. A fully automatic system must include an automatic speech recognition (ASR) module for obtaining automatic transcriptions of the conversations. The acoustic environment on these conversations is unpredictable with a large variety of noise types and intensity. Customers may not be native French speakers and conversations may exhibit frequent disfluencies. The agent may call another service for gathering information. This may cause the introduction of different types of non-speech sounds that have to be identified and discarded. For all these reasons, the word error rate (WER) of the ASR system is highly variable and can be very high. Conversations to be analyzed may be about one or more domain themes. Mentions of relevant themes may be interleaved with mentions of irrelevant comments. Mentions of multiple themes may appear in well separable conversation segments or they may be diffused in zones of a conversation that are difficult to localize. Moreover, mentions can be incomplete or highly imprecise involving repetitions, ambiguities, linguistic and pronunciation errors. Some mentions of an application relevant theme may become irrelevant in certain contexts. For example, a customer may inquiry about an object lost on a transportation mean that was late. In such a case, the loss should be considered as a more relevant theme than the traffic state delay.

In order to properly approach the problems in the above mentioned variety of scenarios and difficulties, methods relying on evaluating the evidence of each feature in the whole conversation should be combined with methods that evaluate feature evidence on specific, automatically detected, conversation segments. In fact, depending on the customer speaking style, mentions of the same pair of themes can be diffused in large portions of a conversation with not clearly defined location for each theme, or they may appear in other conversations in well-localized segments. Thus, it appears

Download English Version:

<https://daneshyari.com/en/article/559011>

Download Persian Version:

<https://daneshyari.com/article/559011>

[Daneshyari.com](https://daneshyari.com)