ACL HLT 2011

**Workshop on Automatic Summarization for
Different Genres, Media and Languages**

Proceedings of the Workshop

23 June, 2011
Portland, Oregon, USA

Introduction

Welcome to the ACL Workshop on Automatic Summarization for Different Genres, Media, and Languages!

Our motivation for organizing this workshop has been the need many researchers in the field have seen, to come together to discuss various new issues that the field is facing as more and more summarization work is being conducted for domains beyond newswire and broadcast news. Extractive summarization of newswire text has dominated summarization research for over a decade. Large corpora of machine- and human-authored summaries have been collected and evaluation has been standardized to a large extent. As work on different genres, media and languages, such as voice mail, email, meetings, broadcast conversations, lectures, chat, blogs and scientific articles becomes more prominent, the need to precisely define tasks, to provide corpora to support comparison between approaches, and to identify desirable evaluation metrics is becoming increasingly urgent. We hope that this workshop will provide a valuable opportunity for all participants to present their work and to engage in discussion about the issues and problems they face, and how we can best support the changing nature of the field.

We have an exciting mix of papers. Some introduce novel summarization tasks: abstractive summarization of line graphs from popular media, summarization of Wikipedia articles with increasing popularity, summarization of chat for the military. Others present new approaches to summarization tasks that have been gaining popularity in recent years, such as summarization of spoken meetings and scientific articles. We are fortunate to have a paper from the organizers of the Text Analysis Conferences (TAC). This paper presents an in-depth analysis of the newest task adopted for the evaluation that, while still based on news, promotes the use of abstractive approaches and makes it possible to track the types of information people consider important and summary-worthy. Finally, we have a reminder that continuity in research focus truly helps to understand a domain and sharpen our understanding of prior approaches to the task; we will hear about some significant improvements in the topic model approach that has proven to be so successful for multi-document summarization of news.

In addition to our diverse program, we will have two invited presentations that will give a more structured overview of the emerging research in summarization. Gabriel Murray's talk is titled "Trends in Abstracting Conversations" and Pascale Fung's is "Structural Summarization of Spoken Documents and its Application to Meeting Minute Generation".

We look forward to a stimulating and engaging workshop!

Ani Nenkova, Julia Hirschberg, and Yang Liu

Organizers:

Ani Nenkova, Univeristy of Pennsylvania
Julia Hirschberg, Columbia University
Yang Liu, University of Texas at Dallas

Program Committee:

Annie Louis (Univ. of Pennsylvania)
Benoit Favre (LIUM, France)
Bonnie Dorr (UMD)
Chin-Yew Lin (Microsoft Research, China)
Daniel Gillick (ICSI)
Dilek Hakkani-Tur (Microsoft)
Fei Liu (UT Dallas)
Gabriel Murray (UBC, Canada)
Gerald Penn (Univ. of Toronto, Canada)
Hakan Ceylan (Yahoo)
Izhak Shafran (Oregon Health and Science University)
Kathy McKeown (Columbia University)
Lucy Vanderwende (Microsoft Research)
Pascale Fung (HKUST, Hong Kong)
Ryan McDonald (Google)
Sameer Maskey (IBM)
Steve Renals (Edinburgh, UK)
Thomas Kleinbauer (DFKI, Germany)
Xiaojun Wan (Peking University, China)

Invited Speakers:

Pascale Fung (HKUST, Hong Kong)
Gabriel Murray (UBC, Canada)

Table of Contents

Workshop Program

09:15-09:30   Introduction

09:30-10:30   Invited talk: Pascale Fung
                        Structural Summarization of Spoken Documents
                        and its Application to Meeting Minute Generation

10:30-11:00   Morning break

11:00-12:30   Poster session

12:30-14:30   Lunch break

14:30-15:30   Invited talk: Gabriel Murray
                        Trends in Abstracting Conversations

15:30-16:00   Afternoon break

16:00-17:30   Discussion

Posters

Plans Toward Automated Chat Summarization
David C. Uthus and David W. Aha

Towards Multi-Document Summarization of Scientific Articles:Making Interesting Comparisons with SciSumm
Nitin Agarwal, Ravi Shankar Reddy, Kiran Gvr and Carolyn Penstein Rosé

Summarizing Decisions in Spoken Meetings
Lu Wang and Claire Cardie

Who wrote What Where: Analyzing the content of human and automatic summaries
Karolina Owczarzak and Hoa Dang

WikiTopics: What is Popular on Wikipedia and Why
Byung Gyu Ahn, Benjamin Van Durme and Chris Callison-Burch

Abstractive Summarization of Line Graphs from Popular Media
Charles Greenbacker, Peng Wu, Sandra Carberry, Kathleen McCoy and Stephanie Elzer

Extractive Multi-Document Summaries Should Explicitly Not Contain Document Specific Content
Rebecca Mason and Eugene Charniak

Posters from other workshops and main conference

Text Specificity and Impact on Quality of News Summaries
    Annie Louis and Ani Nenkova
    Workshop on Monolingual Text-To-Text Generation at ACL-HLT 2011

A Pilot Study of Opinion Summarization in Conversationss
    Dong Wang and Yang Liu
    ACL-HLT 2011 main conference

Why is "SXSW" Trending? Exploring Multiple Text Sources for Twitter Topic Summarization
    Fei Liu, Yang Liu, and Fuliang Weng
    Workshop on Language in Social Media at ACL-HLT 2011