

Research Statement

Naushad UzZaman

University of Rochester

<http://www.cs.rochester.edu/u/naushad>

My research interests are in Natural Language Understanding (NLU) and Natural Language Processing (NLP), with the focus on Temporal Information Processing, Information Extraction, Social Media Text Analysis, Medical NLP, Question Answering, and Multimodal Summarization. In my PhD dissertation, I primarily worked on problems related to Temporal Information Processing that are fundamental to deep language understanding and are key to many applications, such as question answering, information extraction, timeline visualization, and document summarization. These techniques can be applied in news, medical, history and other domains.

Research Context

Artificial Intelligence initially started with the goal of achieving deep language understanding. However, due to the easy access to web-scale data, the NLP community has shifted toward building applications that need a shallow understanding of a large set of documents. Nevertheless, there are other applications that require deep understanding of documents. For example, a doctor wants to know from a patient's historical records, if the patient had any abdominal pain a few months before the CT scan was performed. To answer this question, an automated system needs to understand the medical events, the temporal expressions, the temporal relations, and then make an inference to identify how these events are related to each other in terms of time. This task is not possible to resolve with a shallow understanding of documents.

Another example of how a shallow understanding of language is insufficient can be found in the unrestricted text used in the social media. If we want to build a tool to show prediction tweets for a particular game, then just searching the tweets by matching keywords related to the team-names will not be sufficient. First, we have to understand that it is a *prediction* tweet for a particular game, then we have to identify which team the prediction is for. As a result, just by having a shallow understanding or by matching keywords we cannot build a solution performing all these tasks.

I worked primarily on deep language understanding applications, as the examples mentioned above, in news, medical, social media and dialog system domains. I considered machine learning based solutions or linguistics based solutions or a hybrid between the two, depending on the constraints of the application.

Research Accomplishments

Dissertation Research. My dissertation research focuses on the following areas:

- **Temporal Information Extraction:** We implemented hybrid systems with linguistically motivated solutions and machine learning classifiers for extracting temporal information from raw text. We do deep semantic parsing and use hand-coded rules to extract events, features and temporal expressions from the logical forms produced by the parser. In parallel, we filter events, extract event features, temporal expressions, classify temporal relations using machine-learning classifiers. Our system had a competitive performance in all tasks of TempEval-2 [1] [2] [3].
- **Temporal Evaluation:** We proposed a new metric for evaluation of temporal annotation. Our metric uses temporal closure to reward relations that are equivalent but distinct. It also measures the overall performance of systems with a single score, making comparison between different systems straightforward. Our approach is intuitive and computationally inexpensive. This evaluation metric will be used in the premier temporal annotation shared task TempEval-3 to evaluate participants [4].
- **Temporal Question Answering:** We proposed a temporal QA system that performs temporal reasoning and showed how it can be used to evaluate automated temporal information understanding. With temporal

reasoning, our QA system can answer *list, factoid and yes/no questions*. This was a joint work with Hector Llorens. [5]

- **Creating 1 million word corpus with temporal annotation and organizing TempEval-3:** We teamed up with other participants of TempEval-2, Hector Llorens and Leon Derczynski, and the organizers of TempEval-2, Marc Verhagen and James Pustejovsky, to create a 1 million word corpus with temporal annotation and to organize TempEval-3. We will automatically annotate a 1-million word corpus from news domains by merging our system outputs and human-reviewing a portion of it. We will organize TempEval-3 with that corpus.
- **Merging Temporal Annotations:** We proposed algorithms for merging multiple temporal annotations. With such a merging algorithm, we can merge multiple automated system annotations, multiple human annotations, or a combination of both. We evaluated our algorithm over TimeBank and AQUAINT corpora considering three state-of-the-art systems. The results of merging improve over individual systems and baseline merging algorithms. We will be using this merging tool to merge automated system outputs for TempEval-3 to create the 1-million word corpus. This was a joint work with Hector Llorens [6].

Other Research Highlights.

- **Multimodal Summarization:** In collaboration with Jeffrey P. Bigham and James Allen, I worked on illustrating complex sentences as multimodal summaries combining pictures, simple sentence structure and summarized text. We showed that pictures alone are insufficient to help people understand most sentences, especially for readers who are unfamiliar with the domain. An evaluation of our system in the Wikipedia domain illustrates both the promise and challenge of automatically creating multimodal summaries [7]. MMS could be used to help people with cognitive disabilities, children, older people, or people whose first language is not English [8].
- **Game Prediction with Social Media at Yahoo! Research, Barcelona:** We explored the use of social media for game prediction by prototyping our system on World Cup Soccer 2010 tournament with Twitter data. We extracted around 150K prediction tweets from 1.8 billion tweets by matching the syntactic structure of predictions, instead of just matching prediction words. We created a prediction retrieval exploratory search application with the high-precision extracted predictions, where all predictions for a game are shown and the predictors are ranked according to their previous prediction success. Finally, we aggregated the predictions and we predicted the most likely outcome of upcoming games with high success. A key benefit of our framework is that it does not rely on the domain dependent knowledge for prediction, rather uses crowd's predictions from the social media. As a result, we can consider crowd's analysis about the event and by aggregating the predictions, we can attempt to predict the outcome of other sport events, elections, product release dates and any other future events discussed in the social media. Due to Twitter's restriction on publishing research results with twitter firehose data, eventually this work was not published but I can discuss and present this work.
- **Information Extraction on Demand at Microsoft Medical Media Lab:** I implemented an Information Extraction tool that can learn different patterns from user's examples and extract the pattern instances from natural language texts in medical domain, e.g. release notes, radiology reports, etc. For pattern matching, we used the medical ontology (UMLS) to understand medical patterns like diseases, medicines, etc. and considered language features for general patterns like patient names, IDs, dates, etc. My implemented tool was shared with the clinical staff at Washington Hospital Center and they found it useful.
- **In-Car Dialog System at Bosch RTC:** I developed a syntactic and semantic grammar for a commercial in-car dialog system. It was very similar to the famous *Siri*, iPhone's dialog assistant. My work was in navigation, local business and MP3 domain and my developed grammar could understand sentences like, *"I want to go to Dave's house"*, *"Suggest the nearest Chinese restaurant"*, *"I want to go to the gym on my way home"*, *"Play music by Scorpion"*, *"Raise the volume"*, etc.

- **Data Mining Projects at Bosch RTC:** I worked on a large scale Data Mining project of Breast Cancer Detection. I experimented with different machine learning techniques, e.g. Support Vector Machine (SVM), Decision Tree, KNN, Bayes Net, Neural Network, etc. Finally I built the system using SVM and also implemented the Feature Selection technique to improve the performance and reduce the computation time. Additionally, I also worked on a recommendation system.

Future Research Directions

I will explore the fundamental topics of natural language understanding/processing and apply the techniques in different domains and applications. Following are the brief outlines of several interesting topics:

- **Social Media Text Analysis:** Recently, I was exposed to the social media text analysis and I see lots of opportunities in this domain. My framework for soccer game prediction can be extended to prediction of election, stock market and other events discussed in the social media. I want to explore these possibilities. There are also possibilities for exploratory search applications to display high-precision extractions that I am also interested in. The exploratory search can also help to deliver high quality contextual advertisements, since we know exactly the context of the tweet by extracting information about it.
- **Making information accessible:** The majority of my work falls under extracting information and understanding relations between them. In my work on Multimodal Summarization [7] [8] I worked on making my extracted information accessible. I want to collaborate with Human Computer Interaction colleagues to build more applications to make information accessible to kids, older people, people with disabilities, or even to busy people for quick skimming. My proposed pictorial temporal structure [9] could be one application combining Multimodal Summarization and temporal structure to make information accessible. I want to explore other similar applications.
- **Temporal Summary, Visualization and QA System:** In my dissertation I have done the ground work for building advanced NLP applications with temporal information processing. I want to extend that work to build an application that can give a temporal summary in text and visualization at the same time for a document with the temporal QA capability. This combined application would be very useful for doctors to understand and inquire patient historical records in Medical NLP domain, for students/children to understand and learn history, for people with reading disabilities to follow the news articles, and for many others as well.
- **Temporal Evaluation Shared Task - TempEval:** I have been very active in the TempEval community, first as a participant in TempEval 2010, then as the initiator and co-organizer of TempEval 2013. I would like to continue contributing in that community. In my recent work, I proposed the evaluation of temporal information understanding with temporal question answering [5]. We can easily evaluate systems extracting temporal information in different domains with our released toolkit for temporal QA. I would be interested to initiate and co-organize the next TempEval shared task as well, to evaluate the temporal information understanding.
- **Medical NLP:** There is a push towards converting all medical records to electronic medical records. This will introduce the challenge to access all medical information effectively. It would be very convenient for medical practitioners to have tools to automatically extract information for them. With my experience on building information extraction tool for medical domain combined with my language understanding/processing experience, I want to explore temporal information processing applications and other medical NLP applications to assist medical practitioners.
- **Dialog Systems like Siri:** I have also worked on dialog assistants similar to iPhone's *Siri* for car dialog systems. I want to explore such dialog assistant systems in different domains as well.

I plan to continue my research in natural language processing/understanding in the domain of social media text analysis, medical NLP, news, financial market, history, dialog systems and related domains. I will look for opportunities to collaborate with colleagues in these areas to solve fundamental problems that impact society.

References

- [1] Naushad UzZaman and James F. Allen. “TRIPS and TRIOS System for TempEval-2: Extracting Temporal Information from Text”. *International Workshop on Semantic Evaluations (SemEval-2010), Association for Computational Linguistics (ACL)*, Sweden, July 2010.
- [2] Naushad UzZaman and James F. Allen. “Event and Temporal Expression extraction from raw text: first step towards a temporally aware system”. *International Journal of Semantic Computing*, 2011.
- [3] Naushad UzZaman and James F. Allen. “Extracting Events and Temporal Expressions from Text.” *Fourth IEEE International Conference on Semantic Computing (IEEE ICSC2010)*, Pittsburgh, USA, September 2010.
- [4] Naushad UzZaman and James F. Allen. “Temporal Evaluation”. *Proc. of The 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (Short Paper)*, Portland, Oregon, USA, June 2011.
- [5] Naushad UzZaman, James F. Allen and Hector Llorens. “Evaluation of Temporal Information”. *In Progress*.
- [6] Naushad UzZaman, Hector Llorens and James F. Allen. “Merging Temporal Annotations”. *Submitted*.
- [7] Naushad UzZaman, Jeffrey P. Bigham and James F. Allen. “Multimodal Summarization of Complex Sentence”. *Proc. of International Conference on Intelligent User Interfaces (IUI)*, Palo Alto, California, 2011.
- [8] Naushad UzZaman, Jeffrey P. Bigham and James F. Allen. “Multimodal Summarization for people with cognitive disabilities in reading, linguistic and verbal comprehension”. *2010 Coleman Institute Conference*, Denver, CO, 2010.
- [9] Naushad UzZaman, Jeffrey P. Bigham and James F. Allen. Pictorial Temporal Structure of Documents to Help People who have Trouble Reading or Understanding. *International Workshop on Design to Read, ACM Conference on Human Factors in Computing Systems (CHI)*, Atlanta, GA, April 2010.