IBM Cognitive Computing - An NLP Renaissance!

Salim Roukos
Senior Manager of Multi-Lingual NLP and CTO for Translation Technologies
IBM T. J. Watson Research Center
roukos@us.ibm.com

Abstract

Electronically available multi-modal data (primarily text and meta-data) is unprecedented in terms of its volume, variety, velocity, (and veracity). The increased interest and investment in cognitive computing for building systems and solutions that enable and support richer human-machine interactions presents a unique opportunity for novel statistical models for natural language processing.

In this talk, I will describe a journey at IBM during the past three decades in developing novel statistical models for NLP covering statistical parsing, machine translation, and question-answering systems. Along with a discussion of some of the recent successes, I will discuss some difficult challenges that need to be addressed to achieve more effective cognitive systems and applications.

About the Speaker

Salim Roukos is Senior Manager of Multi-Lingual NLP and CTO for Translation Technologies at IBM T. J. Watson Research Center. Dr. Roukos received his B.E. from the American University of Beirut, in 1976, his M.Sc. and Ph.D. from the University of Florida, in 1978 and 1980, respectively. He joined Bolt Beranek and Newman from 1980 through 1989, where he was a Senior Scientist in charge of projects in speech compression, time scale modification, speaker identification, word spotting, and spoken language understanding. He was an Adjunct Professor at Boston University in 1988 before joining IBM in 1989. Dr. Roukos has served as Chair of the IEEE Digital Signal Processing Committee in 1988.

Salim Roukos currently leads a group at IBM T.J. Watson research Center that focuses on various problems using machine learning techniques for natural language processing. The group pioneered many of the statistical methods for NLP from statistical parsing, to natural language understanding, to statistical machine translation and machine translation evaluation metrics (BLEU metric). Roukos has over a 150 publications in the speech and language areas and over two dozen patents. Roukos was the lead of the group which introduced the first commercial statistical language understanding system for conversational telephony systems (IBM ViaVoice Telephony) in 2000 and the first statistical machine translation product for Arabic-English translation in 2003. He has recently lead the effort to create IBM's offering of IBM Real-Time Translation Services (RTTS) a platform for enabling real-time translation applications such as multilingual chat and on-demand document translation.

Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), page 1,
October 25-29, 2014, Doha, Qatar.