ISCA DL Tour Report in Santiago (Oct. 6-7) and Belem (Oct. 9-10)
Prepared by Chin-Hui Lee (chl@ece.gatech.edu)

(A) Santiago, Chile, Oct. 6-7, 2008

(i) Names of the local hosts
Néstor Becerra Yoma, PhD
Speech and Processing Laboratory
Department of Electrical Engineering
Universidad de Chile

Av. Tupper 412-3, POBox: 412-3
Santiago-CHILE

(ii) Lecture 1: on Oct. 6, 2008 at 10:30-12:00 at University of Chile

Teaching Computers to Speak and Listen: A Wonderful Journey from Science Fictions to Technological Realities

Making machines to speak human languages and respond to voice commands has always been a challenge for scientists and engineers in modern history. Although such a dream is not easily realized in real life, it is very common to see machines, computers and droids speak and listen in science fictions. Some of such famous machines include HAL in *2001: A Space Odyssey*, and R2D2 and C3PO in *Star Wars*. In the real world speaking machines were first demonstrated by Homer Dudley of Bell Laboratories in the 1939 World Fair in New York City. The technology has come a long way that we can now design machines to speak fluently multiple languages on various topics. On the other hand primitive listening machines were first developed at Bell Labs in the 1940s to understand English digits. Nowadays we have witnessed many automated services and products that take human voice as inputs. We can even purchase voice typewriters at a very reasonable price. In this lecture we will go over the wonderful journey of developing machines to listen and speak, and illustrate their current capabilities and limitations. It is clear speech and language research is still needed to advance today's technologies so we can realize some of the promises issued in science fictions for computers to be collaborative voice communicators to help human beings solve problems.

(iii) Lecture 2: on Oct. 7, 2008 at 10:30-12:00 at University of Chile

A Critical Overview on Model Adaptation in Speech, Language and Media Processing

Human brains adapt to changing situations by combing previous experience in individual's lives and the new conditions they are exposed to. This scenario can be simulated with Bayesian learning well known in statistics literatures. Although the human perception and adaptation process is not exactly known, the mathematical formulation of Bayesian learning has been adopted in many practical data mining processing, including speaker model adaptation to a new speaker in automatic speech and speaker recognition, language adaptation in multilingual speech recognition, language model adaptation to unknown task domains, and concept model adaptation in automatic image annotation. In this lecture we present the mathematical foundations of Bayesian learning and provide a comprehensive overview of the technology advances in automatic speech recognition. We then discuss recent development of adaptation algorithms in speech, language and media processing. Finally we address some remaining challenges in bridging the gap between model adaptation of human beings and machines. (Reference, Lee and Huo, *Proc. IEEE*, August 2000)

(iv) Attendance: Only local students in Electrical Engineering attended the lectures, no faculty members showed up. I guess the promotion can be done better by the local host.

(v) Discussion: I had productive face-to-face discussions with many of Prof. Becerra Yoma's students. Most of them are eager to learn. We also discuss how to promote speech and language research and development in Chile.

(vi) Expected impacts: Prof. Becerra Yoma expressed interest in hosting an ISCA ITRW on language learning as a first step to bring the South American speech community closer to the international counterpart. As a follow-up he will also consider bidding for a future Interspeech conference in Santiago, Chile.

(vii) Side comments and concerns: I believe the local promotion should be done better. The local attendance was not good in Santiago, possibly because there is no local industry and other speech research groups in Santiago. The DL Tour can be more effective and wider spread if it can be coupled with a local workshop like what was done last year during my DL tour in Bangkok, Thailand in November 2007.

(B) Belem, Brazil, Oct. 9-10, 2008

(i) Names of the local hosts
Aldebaro Barreto da Rocha Klautau Junior
Federal University of Para (Universidade Federal do Para, UFPA)
Electrical and Computer Engineering Department

Phone: +55 (91) 3201-7674
UFPA - PPGEE - LaPS

CP: 8619 - CEP: 66075.110
Belem - Para - Brazil

(ii) Lecture 1: on Oct. 9, 2008 at 15:00-16:30 at Federal University of Para

Teaching Computers to Speak and Listen: A Wonderful Journey from Science Fictions to Technological Realities

Making machines to speak human languages and respond to voice commands has always been a challenge for scientists and engineers in modern history. Although such a dream is not easily realized in real life, it is very common to see machines, computers and droids speak and listen in science fictions. Some of such famous machines include HAL in *2001: A Space Odyssey*, and R2D2 and C3PO in *Star Wars*. In the real world speaking machines were first demonstrated by Homer Dudley of Bell Laboratories in the 1939 World Fair in New York City. The technology has come a long way that we can now design machines to speak fluently multiple languages on various topics. On the other hand primitive listening machines were first developed at Bell Labs in the 1940s to understand English digits. Nowadays we have witnessed many automated services and products that take human voice as inputs. We can even purchase voice typewriters at a very reasonable price. In this lecture we will go over the wonderful journey of developing machines to listen and speak, and illustrate their current capabilities and limitations. It is clear speech and language research is still needed to advance today's technologies so we can realize some of the promises issued in science fictions for computers to be collaborative voice communicators to help human beings solve problems.

(iii) Lecture 2: on Oct. 10, 2008 at 10:30-12:00 at Federal University of Para

Statistical Natural Language Processing: Lessons from Automatic Speech Recognition

Many of the recent advances in statistical natural language processing (NLP) are mainly attributed to a number of key factors, namely: (1) Shannon's source-channel characterization of NLP problems; (2) the availability of large collections of labelled text data; (3) the implementation of efficient and effective machine learning algorithms; and (4) increase in computing power to learn statistical properties in data and real-time execution of may real-world NLP applications. This same data-driven, pattern matching paradigm prevails in the field of automatic speech recognition (ASR). We review the fundamentals in ASR and extend the same paradigm to some popular NLP tasks, such as *N*-gram language modeling, part of speech tagging, text understanding, statistical parsing, text categorization, and machine translation. The learning curve in ASR development in the last thirty years gives up plenty of lessons to predict what's laying ahead of us in research and applications in statistical NLP.
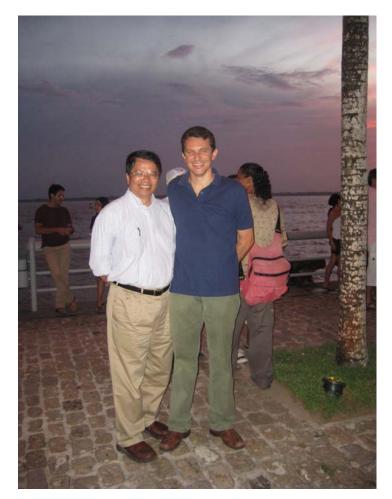
(iv) Attendance: Many undergraduate and graduate students attended both lectures. Quite a few faculty members showed up as well. The whole conference room was packed. I was very impressed with the enthusiasm expressed. The local host did a very good job.

(v) Discussion: I had productive face-to-face discussions with many students. Most of them are eager to learn. I had also had interesting discussions with faculty members.

(vi) Expected impacts: Prof. Aldebaro Klautau expressed interest in expanding his speech research team, and will increase his interactions with the international community. A few of his students had internship and visiting scholar experience in the US and EU countries. There is also a small industry interested in speech applications.

(vii) Side comments and concerns: I believe the local promotion is excellent in this case. The DL Tour can be more effective and wider spread if the invitations are extended to both undergraduate

and graduate students in the host institute and other local groups like what was done last year during my DL tour in Manila, Philippines in November 2007. Just like in Manila a local reporter interviewed me after my lectures. A news clip will be sent to the ISCA Board when it is available,

(C) Summary

(i)   Overall it has been a great experience for me to interact with faculties, students, researchers and engineers from resource-limited countries. I think ISCA is doing a good service to the community by providing such services. Many of my hosts expressed the desire of having similar activities in the future because it is difficult for them to have prominent visitors, and it is much cheaper to bring one lectures to a large local audience than to send researchers out to conferences.

(ii)   The local hosts play a key role in the success of a DL visit. Some requirements can be imposed on local promotion because a visit is approved. In my case, Santiago is the only city that a better promotion can be beneficial to the local community. The other three cases are very well-executed.

(iii) Some care is needed when visiting countries and cities that language communication can be a problem. Local safety can also be a major concern.

(iv)   An ISCA process can be established. I had problems in having my reimbursement issued. There was no memory in the process so that the same mistake was made for my case in 2008 even we had gone through the same steps in 2007. A single contact person or officer is needed. A better efficiency is mandatory. I had a terrible time contacting ISCA, and in some case missing a window of getting low-cost airfares.

(v)   Visa expenses can be substantial and need to be considered: Chile ($131), Brazil ($130). Requirement needs to be checked early. I almost missed applying for Brazil.