

Who should attend

The Workshop is aimed at Teachers, Researchers and Industry personnel who want to move into the Speech Recognition area, with motivation for building practical systems and applications using available tools and resources. They should have an aptitude to do hands-on work for building working systems even with minimum functionalities and performance – aspects that can then be incrementally improved through further work in an ongoing manner. They are also expected to have a minimum proficiency in computer programming and some exposure to Linux commands. Also recommended strongly is a prior reading of the speech tutorial <http://speech.tifr.res.in/chief/>

How to Apply

Details of Application format, Registration charges, etc are available at:

<http://www.au-kbc.org/speech/>

Admission based on first come first served.

Number limited due to the need for laboratory facilities - Apply before 20th Jan 2011

Contact

For any further information, contact the Workshop Coordinator -

Dr. G. Bharadwaja Kumar,

Member Research Staff,

AU-KBC Research Centre,

MIT Campus of Anna University,

Chromepet, Chennai -600044

Telephone: (+91) 044 2223 2711 extn: 130

Fax: (+91) 044 2223 1034

e-mail: speechgroup@au-kbc.org

URL: www.au-kbc.org

About AU-KBC Research Centre

The Anna University – K B Chandrasekhar Research Centre in the MIT Campus of Anna University Chennai carries out applied research in a unique Public-Private-Partnership (PPP) mode to generate Intellectual Property (IP) of social and economic value to India. Working with in an Innovation Ecosystem, the AU-KBC Centre aims to become fully self-supporting through the commercialization of its IP with in transparent framework of benefit sharing with all the stake holders. Working in the domains of Life Sciences and Information Sciences, the AU-KBC Centre has been active in the area of Language Technology Development for Indian languages, and has recently started working on Speech Application Development.

Automatic Speech Recognition using *Sphinx* and *HTK* – A hands-on Workshop

18-20 Feb 2011

AU-KBC Research Centre
(www.au-kbc.org)



Anna University

Madras Institute of Technology Campus
Chromepet
Chennai – 600044

Funded by:

Department of Science and Technology,
LDC-IL, CIIL, Mysore.
Department of Information Technology

The Background of the Workshop

With the ever decreasing cost of computing and communication, man-machine interfaces are becoming more 'natural', and speech input-output systems are a key driver of this trend. Automatic Speech Recognition (ASR) systems that convert speech in a language to valid text sentences in that language is one of the key components of this package. While even uneducated humans understand speech effortlessly and carry on highly nuanced and subtle conversation with ease using speech, making computers do the same has proven to be an extremely challenging task – both in terms of our understanding of what is involved as well as our ability to implement them.

Building a robust ASR System is highly complex due to the large number of variables involved - type of speaker, type of speech, size of vocabulary, environment in which speech is produced, medium through which speech has passed, etc. While a robust design should accommodate the variabilities in each of these factors, satisfactory systems have so far been built only under restricted contexts and for limited purposes. Even this is limited to English, Japanese and some European languages; ASR systems of even moderate performance are yet to be built for Indian languages.

The Objectives of the Workshop

The main objective of this workshop is to introduce young faculty and researchers to the area of ASR, with particular attention paid to Indian language ASR research. The Faculty will provide the minimum background knowledge necessary to understand the models, approaches and technologies of ASR. The workshop would focus on the practical aspects of building ASR systems using available software tools so that the participants would acquire some knowledge and skills to build on after the workshop is over.

About Sphinx and HTK

Sphinx and HTK are Free/Open Source Software Tools for ASR research and development work. *Sphinx* is a flexible, modular and pluggable ASR framework to enable new innovations in speech recognition system design. It was created collaboratively by a number of research groups such as the CMU, Sun Microsystems, MERL, HP, UCSC and MIT, and released to the public under the GPL license.

HTK (Hidden Markov Model Toolkit) is another most prominent free software tool for building speech recognition applications. HTK was originally developed at the Machine Intelligence Laboratory of the Cambridge University, and subsequently developed further through collaborative work. HTK license also allows certain form of the free use of the source code. Both Sphinx and HTK provide state of the art frameworks and technologies for building ASR applications of practical value, and are a boon to the ASR research community.

The Workshop Contents and Coverage

The topics covered through lectures and lab classes include:

- Introduction to ASR concepts and principles
- Acoustic and Language Modeling
- ASR in landline/cellular environments
- Speech and Speaker Adaptation
- Language Identification
- Natural Language Processing in ASR
- ASR with Sphinx and HTK
- ASR in Indian Languages.

About 70% of the workshop will consist of lab classes using Sphinx and HTK

The Workshop Faculty

The workshop faculty are active researchers and experts in the speech/language technology area from leading institutions in India.

- Dr. G. Bharadwaja Kumar, AU-KBC Centre
- Dr. K. Samudravijaya, TIFR, Mumbai
- Prof. K. Narayana Murthy, CIIL, Mysore
- Prof. S. Umesh, IIT Madras, Chennai
- Dr. L. Sobha, AU-KBC Centre
- Dr. T. Nagarajan, SSN College, Chennai
- Dr. A. Shahina, SSN College, Chennai
- Mr. Vikas Joshi, IIT Madras, Chennai
- Experts from Industry