

Keynotes

Speech translation for everyone - breaking down the barriers

Arul Menezes, Microsoft Research

Abstract

Fifty years ago Star Trek had the Universal Translator. Thirty-five years ago we were introduced to the babel fish in The Hitchhiker's Guide to the Galaxy. Decades later, is reality finally catching up to science fiction? Given the enormous strides made in speech recognition and machine translation over the last decade, is this just a matter of chaining speech recognition and machine translation together?

In the Skype Translator project we set ourselves an ambitious goal - to enable successful open-domain conversations between Skype users in different parts of the world, speaking different languages. As one might imagine, putting together two error-prone technologies such as speech recognition and machine translation raises some unique challenges.

In this talk, I will share what we have learned over the course of the Skype Translator project. I will discuss what we are doing to bridge the gap between ASR and MT, how we are adapting our ASR and MT systems to the real world challenges presented by our open-domain conversational scenario, and what it takes to get this technology into the hands of real users. I will also touch upon some of the open issues and challenges we still face.

Bio

Arul Menezes heads the Machine Translation team at Microsoft Research. Over the past 15 years, he has driven Machine Translation at Microsoft Research from a basic research project to a web-scale production service with a variety of offerings for consumers and businesses, and millions of users worldwide. These include the Bing Translator and the Microsoft Translator Hub customization service, as well as the upcoming Skype Translator product. The Microsoft MT system is based on the treelet translation approach to syntactic statistical MT, co-invented by Arul, Chris Quirk and Colin Cherry. The MSR MT team integrates research and product development in a single team, covering everything from MT modelling and algorithms to data gathering and delivery of the live web service. This eliminates the traditional "tech transfer" from research to product, and enables the team to get research breakthroughs into customer hands without delay. Arul was educated at the Indian Institute of Technology, Bombay and at Stanford University.