Indexing WFST for Spoken Term Detection

Lucas Ondel (iondel@fit.vutbr.cz)

Thanks to many recent innovations, the availibity of mutlimedia documents has increased and it has raised the need for solutions to efficiently index and search these kind of archives. Speech Retrieval (SR) is a technology which integrates automatic speech recognition (ASR) and information retrieval, two technologies combined together to provide large scale access to spoken content. However, this is not a straigthforward task, as state-of-the-art ASR systems are not able to transcribed reliably unconstraint speech and considering the heterogeneous nature of the large spoken databases, SR is not simply indexing the text output of an ASR system. Thus it raises to main problem : how to represent the uncertainty of the ouput of the ASR system, and how to create our index according to this uncertainty?

The presentation will talk about this two problems and we draw here its global organisation. First of all, we will introduce Weighted Finite State Transducers (WFST), their mathematical definition and a few basic operations to work with them. Then, we will study how we can answer to our first problem with the WFSTs by representing the ASR output by a WFST (a so-called WFST) and last, we will be able to see how we can index these WFSTs and find a solution to the last problem.