

Conclusiones y trabajo a futuro

El objetivo de este proyecto era realizar una evaluación de un reconocedor fonético para el español hablado en México. En parte fue motivado por el hecho de que en México no existe ninguna investigación previa sobre el tema. Este trabajo puede servir como punto de partida para futuras investigaciones en reconocimiento fonético del lenguaje español.

Durante el desarrollo de este proyecto, se hizo un experimento base, el cual consistió en crear un reconocedor independiente del contexto. Sobre este experimento base se realizaron una serie de mejoras, las cuales consistieron en:

- La reducción de errores de inserción ajustando los límites de duración.
- La incorporación de modelos dependientes del contexto.
- El entrenamiento del reconocedor con una mayor cantidad de muestras por categoría.

Después de estas mejoras la evaluación del reconocedor mostró 55.3% sobre los datos de evaluación.

La prueba final se realizó sobre habla espontánea (*stories*), obteniendo un desempeño de 47.7% . La reducción en el desempeño se debió principalmente a dos razones: primero, los datos del subconjunto de prueba eran desconocidos para el reconocedor; segundo, *stories* es un conjunto de datos difícil de reconocer por su espontaneidad, por las variaciones en el ritmo y entonación del habla y por la cantidad de disfluencias (em, ah, este, etc.) contenida en ellos.

Uno de los factores que influye en el desempeño del reconocedor es la calidad del corpus y las condiciones de grabación. Se observó que ciertos locutores en el corpus obtenían muy bajo nivel de reconocimiento, ya que durante la grabación existía mucho ruido de fondo o en la línea telefónica. A pesar de ello se dejaron, porque de esta manera el reconocedor se “acostumbra” a diferentes tipos de ambiente, donde las condiciones no son óptimas.

Una dificultad que surgió durante la realización de este proyecto fue la parte de evaluación, donde la herramienta de evaluación del CSLU Toolkit no estaba trabajando de forma correcta, penalizando desempeño en situaciones donde realmente no era tan grave el error. En su lugar, para hacer una justa evaluación del desempeño del reconocedor, se realizó la evaluación del sistema utilizando el software de evaluación NIST. Finalmente, aunque el nivel de reconocimiento pareciera ser bajo, los resultados se comparan favorablemente con investigaciones hechas en inglés.

Una de las motivaciones para esta tesis era que si se mejora el reconocimiento fonético, es muy probable que se mejore el reconocimiento de palabras. Un ejemplo de ello es el desempeño mostrado por el reconocedor en el dominio de dígitos.

Al final de este proyecto, se cuenta con una evaluación de reconocimiento a nivel de fonemas para el español hablado en México. Los resultados y el corpus de estos experimentos se encuentran publicados en Internet en la dirección: http://info.pue.udlap.mx/~sistemas/tlatoa/rec_fonetico.html. Este material está disponible para cualquier grupo interesado en hacer investigación en esta área y así podrán comparar directamente sus resultados y mejoras.

Entre los trabajos a futuro están los siguientes:

- Corregir errores cometidos en la división de las clases generales.
- Mejorar la clasificación de fonemas.
- Experimentar con el límite máximo de duraciones.
- Mejorar el desempeño utilizando modelos probabilísticos de lenguaje natural.
- Experimentar con los números máximos de muestras por categoría para incluir todos los datos de fonemas.
- Agregar los *stories* que no se utilizaron a los datos de entrenamiento y evaluación, para tener datos más similares en los tres conjuntos.
- Realizar un análisis más profundo de errores.