

CALL FOR PARTICIPATION

The Blizzard Challenge 2010

Evaluating corpus-based speech synthesis on common databases

Alan W Black, Simon King, and Keiichi Tokuda

January 31st 2010

In order to better understand and compare research techniques in building corpus-based speech synthesizers on the same data, the annual Blizzard Challenges 2005–2009 were held. We are now pleased to call for participation in the sixth challenge: Blizzard Challenge 2010.

The basic challenge is to take the released speech data, build synthetic voices, and synthesize a prescribed set of test sentences. The sentences from each synthesizer are then evaluated through extensive listening tests. The 2010 Challenge will use a similar a hub and spoke design to the 2009 Challenge. The tasks for 2010 are as follows:

- British English:
 - Hub task EH1: 5 hour (4000 utterance) ‘rjs’ database (male professional speaker, RP accent) kindly supplied by Phonetic Arts and available at 16kHz and 48kHz sampling rates, along with standard Festival labels produced by the University of Edinburgh
 - Hub task EH2: 1 hour (1000 utterance) subset of the ‘Roger’ database (male speaker, RP accent, ARCTIC sentences, same as in 2009) from the University of Edinburgh, with new hand-corrected labels kindly supplied by iFLYTEK, along with standard Festival labels produced by the University of Edinburgh
 - Spoke task ES1: 100 sentences from the ‘Roger’ ARCTIC set with standard Festival labels
 - Spoke task ES2: speech in noise task, with the evaluation focusing on intelligibility, using the same data as EH1
 - Spoke task ES3: higher sampling rate (48kHz) voice, using the same data as EH1
- Mandarin Chinese
 - Hub task MH1: 9 hour (6000 utterance) database kindly supplied by the Chinese Academy of Sciences (6.5 hours as used in 2008 plus 2.5 hours of new data)
 - Hub task MH2: 1 hour (800 utterance) subset of the new portion of the MH1 data
 - Spoke task MS1: 100 sentences from the new portion of the MH1 data
 - Spoke task MS2: speech in noise task, with the evaluation focusing on intelligibility, using the same data as MH1

For all tasks except ES3, 16bit/16kHz wav files should be submitted.

Blizzard Workshop: The results will be presented at a workshop, which will be a satellite of the Seventh ISCA Tutorial and Research Workshop on Speech Synthesis (SSW7) in Kyoto, Japan (ssw7.org). The workshop is scheduled for Sept 25th 2010. A condition of participation in the Blizzard Challenge is the submission of a paper to this workshop. Accepted papers will be presented at the workshop, and published on the Blizzard website.

Website: http://www.synsig.org/index.php/Blizzard_Challenge_2010

Registration: Interested parties should register as soon as possible, by mailing blizzard@festvox.org. They should identify a contact person in their team, and provide email and postal addresses. A registration fee of USD 750 will be required to partially meet the costs of running the challenge and paying listeners. Payment method details will be provided to participants.

Licenses: the licenses for the released data can be found via the Blizzard website. Data will be released to each participant once the appropriate license(s) have been agreed to, but no sooner than Monday 1st February 2010. The procedures for this are given on the Blizzard website.

Mailing list: please join the blizzard-discuss mailing list, which is used to send important information to participants. Instructions are on the Blizzard website.

Timeline

Feb	1	2010	Database release
Mar	29	2010	Test sentences released
Apr	9	2010	Deadline for submitting synthesized speech
Apr	19	2010	Evaluation system goes live
June	11	2010	End of evaluation
June	25	2010	Results distributed to teams
July	23	2010	Satellite workshop papers due
Aug	13	2010	Notification of acceptance
Sep	22-24	2010	SSW7, Kyoto, Japan
Sep	25	2010	Blizzard Challenge Workshop, Kyoto, Japan
Sep	26-30	2010	Interspeech 2010, Makuhari, Japan

Further Information: For further information please contact blizzard@festvox.org.

A description of the challenge was published at Interspeech 2005:
<http://www.festvox.org/blizzard/bc2005/IS051946.PDF>