

The Blizzard Challenge 2011

Evaluating corpus-based speech synthesis on common databases

Simon King, Alan W Black, and Keiichi Tokuda

March 4th 2011

In order to better understand and compare research techniques in building corpus-based speech synthesizers on the same data, the annual Blizzard Challenges 2005–2010 were held. We are now pleased to call for participation in the seventh challenge: Blizzard Challenge 2011. The challenge is to take the released speech data, build one or two synthetic voices, and synthesize a prescribed set of test sentences. The sentences from each synthesizer are then evaluated through extensive listening tests. The tasks for 2011 are as follows:

- American English:
 - Hub task EH1: a general-purpose voice from the full ~ 12000 utterance ‘Nancy’ data set (female professional speaker, US accent) kindly supplied by Lessac Technologies Inc., and available divided into utterances at 16kHz sampling rate, or as original studio session files at 44.1kHz and 96kHz sampling rates.
 - Spoke task ES1: a voice from the full data set designed to read names and addresses (US format)

For both tasks, 16bit wav files at any sampling rate can be submitted. The main listening tests will use wav files downsampled to 16kHz. If there are sufficient entries at higher sampling rates, an additional listening test will be conducted.

Please note that the Blizzard Challenge 2011 runs concurrently with Phase One of the Blizzard Challenge 2012. We encourage participation in both challenges. For this reason, the range of tasks for 2011 is smaller than in some recent challenges.

Blizzard Workshop: The results will be presented at a workshop, which will be a satellite of Interspeech 2011, Florence, Italy (interspeech2011.org). The workshop will be scheduled for just before or just after Interspeech, at a location in Italy. 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_2011

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 GBP 500 will be required to partially meet the costs of running the challenge and paying listeners.

License: the license for the released data can be found via 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

Mar	4	2011	Databases released
Apr	15	2011	Test sentences released
Apr	22	2011	Deadline for submitting synthesized speech
Apr	29	2011	Evaluation system goes live
Jun	25	2011	End of Evaluation
Jul	4	2011	Deadline for returning the participant questionnaire
Jul	8	2011	Results distributed to teams who have returned their questionnaire
Jul	22	2011	Satellite workshop papers due
Jul	29	2011	Notification of acceptance
Aug/Sep		2011	Presentation of results at the satellite workshop in Italy
Aug	28-31	2011	Interspeech 2011, Florence, Italy

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>