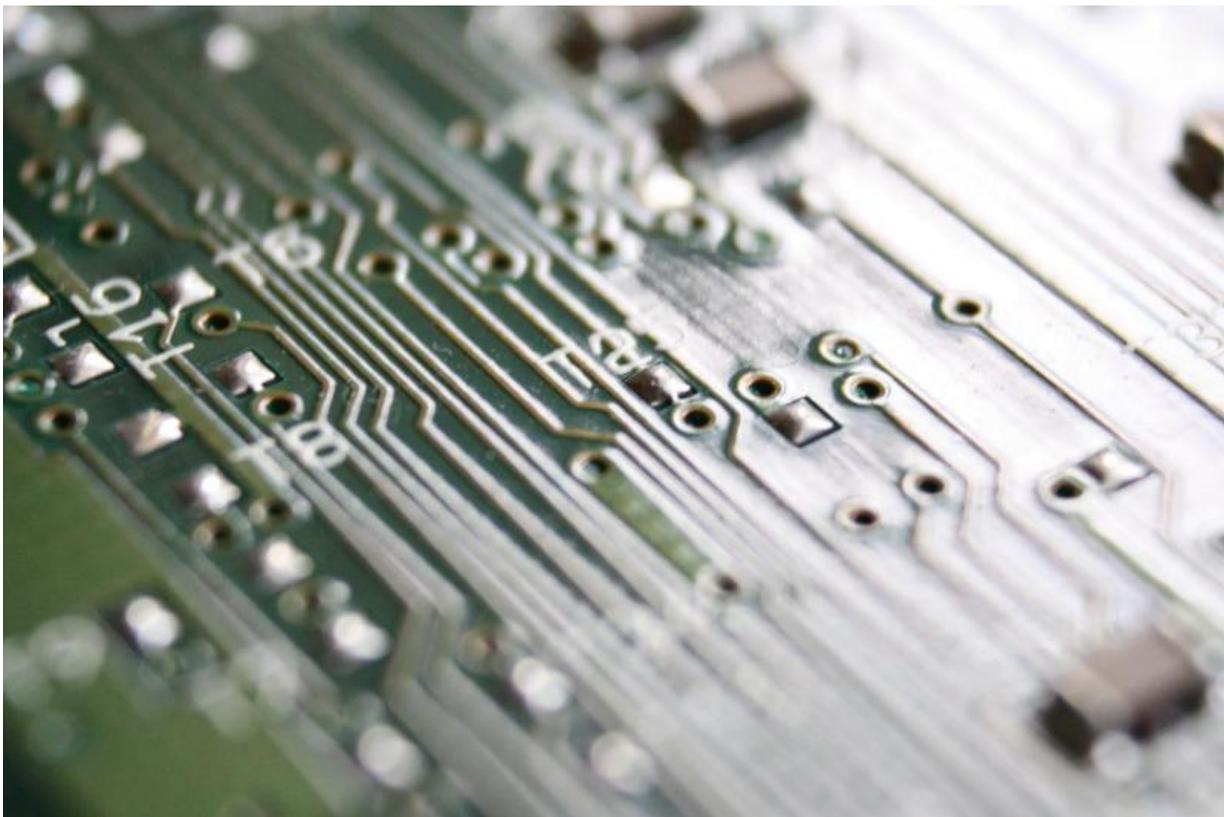


Machine-written speech reads like real political spiel

January 26 2016, by Nancy Owano



Credit: Public Domain

It's high time to stand up and be counted for...We need a true leader who will...Let's bring back honor and....Plain folks like you deserve to be... Issues and platforms aside, political campaign speeches dull the senses at

some point when all candidates on all platforms begin to sound alike.

Similar to the word "lawyer," the word "politician" has taken on a cynical meaning, suggesting someone who says only what will garner applause.

This is not a recent credibility problem, either. Nikita Khrushchev is credited with the quote, "Politicians are the same all over. They promise to build a bridge even where there is no river." Translated from the Portuguese, José Maria de Eça de Queiroz is credited with having said, "Politicians and diapers should be changed frequently and all for the same reason."

MIT Technology Review observed how U.S. congressional floor debates are numerous and "are also remarkably similar. These speeches tend to follow a standard format, repeat similar arguments, and even use the same phrases to indicate a particular political affiliation or opinion. It's almost as if there is some kind of algorithm that determines their [content](#)."

Indeed. The sameness of characteristics and structures in political speeches is what has interested Valentin Kassarnig, College of Information and Computer Sciences at the University of Massachusetts Amherst. So much so that he has written a paper posted on the arXiv this month, titled "Political Speech Generation."

The question that Kassarnig posed in his research: Can a machine generate a speech where you cannot tell the difference between it and a speech that was hand-written? "Many political speeches show the same structures and same characteristics regardless of the actual topic. Some phrases and arguments appear again and again and indicate a certain political affiliation or opinion. We want to use these remarkable patterns to train a system that generates new speeches," said the author.

Kassarnig performed an experiment where he created an [artificial intelligence](#) machine that learned how to write political speeches similar to real speeches and discussed the results.

The author noted past research in the field of Natural Language Generation (NLG) where a collection of example inputs is mapped to output texts of the corpus. "This is basically what we plan to do because we have already all the speech segments labeled with the political party and the opinion. However, our generator will have a simpler architecture."

He discussed the system SciGen, an automatic computer science research paper generator developed by three MIT students. The random papers it creates "show actually a very high quality in terms of structuring and lexicalization, and they even include graphs, figures, and citations."

As for his data source, the main source was the Convote data set, which contained 3857 speech segments from 53 US Congressional floor debates from the year 2005.

Results? "In an experimental evaluation our system performed very well. In particular, the grammatical correctness and the sentence transitions of most speeches were very good. However, there are no comparable systems which would allow a direct comparison."

Based on the author's comments, however, it is not likely that you will see him churning out text for vote-seekers any time soon. "Despite the good results it is very unlikely that these methods will be actually used to generate speeches for politicians," he wrote.

Rather, he said, "the approach applies to the generation of all kind of texts given a suitable dataset. With some modifications it would be

possible to use the system to summarize texts about the same topic from different source, for example when several newspapers report about the same event."

More information: Political Speech Generation, arXiv:1601.03313
[cs.CL] arxiv.org/abs/1601.03313

Abstract

In this report we present a system that can generate political speeches for a desired political party. Furthermore, the system allows to specify whether a speech should hold a supportive or opposing opinion. The system relies on a combination of several state-of-the-art NLP methods which are discussed in this report. These include n-grams, Justeson & Katz POS tag filter, recurrent neural networks, and latent Dirichlet allocation. Sequences of words are generated based on probabilities obtained from two underlying models: A language model takes care of the grammatical correctness while a topic model aims for textual consistency. Both models were trained on the Convote dataset which contains transcripts from US congressional floor debates. Furthermore, we present a manual and an automated approach to evaluate the quality of generated speeches. In an experimental evaluation generated speeches have shown very high quality in terms of grammatical correctness and sentence transitions.

© 2016 Tech Xplore

Citation: Machine-written speech reads like real political spiel (2016, January 26) retrieved 20 September 2024 from
<https://techxplore.com/news/2016-01-machine-written-speech-real-political-spiel.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.