

Partisan Responses Generated from Local Knowledge Graphs

Alexandra Butoi, Jannik Gut, Mian Zhong

Department of Computer Science, ETH Zürich

Email: abutoi@student.ethz.ch, jgut@student.ethz.ch, mzhong@student.ethz.ch

Abstract—Although current Natural Language Processing(NLP) methods excel at extractive question answering (often on datasets specifically designed for this task, such as SQuAD) and text generation, producing high-quality text based on information stored in knowledge graphs still remains an open task. In this paper we apply two models to the U.S. Congressional Record corpus to generate text, namely GPT-2 and GraphWriter. We compare the performance of these current NLP methods in open-domain question answering and analyse the reasons why these architectures fail to produce satisfactory results, identifying some areas that still need further advances. We notice that the text produced using GPT-2 is grammatically correct and coherent but it lacks background knowledge. As it is not designed for question answering, it merely outputs some text after a prompt (the question), which is generally on the same topic but often it does not specifically answer the question. GraphWriter attempts to solve this issue by incorporating information from knowledge graphs into the generated text. This graph Transformer has drawbacks as well, as it was originally designed for a very specific domain with limited types of data. Altering its architecture for open-domain question answering proves to be a difficult task. The output shows potential for opinionated text generation; however, further investigation is needed in order to produce high-quality text.

I. INTRODUCTION

Question Answering (QA) is one of the main tasks in Natural Language Processing (NLP). The goal is to provide an answer to a question posed in natural/spoken language, often from a structured source of information such as a knowledge graph (see Figure 1). A large proportion of the work in this area ([1], [2], [3], [4]) focuses on extractive question answering (i.e. given a question and a paragraph, identify the answer to the question in the paragraph). Even though the reading comprehension performance of such systems has increased considerably since the advent of Transformer-based neural network architectures[5] which can produce very accurate answers, such models still come with several shortcomings. The user of such a system is provided only with a span of text from the given context that contains the answer, not a full sentence or paragraph. Also, if someone wants to search for the answer in a collection of documents, a potential answer will be returned for each document. Merging these answers into one coherent response requires further work. Recently, there has been an increased interest in text generation from knowledge graphs ([6], [7], [8]). This work focuses on providing the language models with external knowledge in the attempt to generate text that is

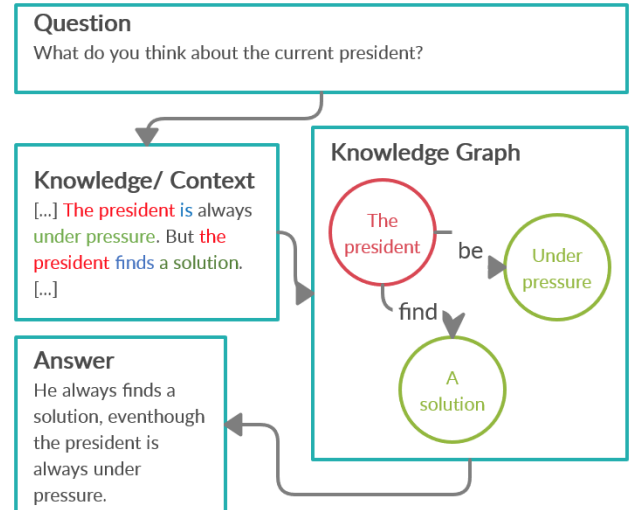


Figure 1. Pipeline to generate a response via knowledge graphs.

factually correct. However, such methods still have to be adapted to the field of question answering.

We adjust the model from Koncel-Kedziorski et al.[8] for a dataset in the political domain with human-curated questions and speeches from the U.S. Congressional Record. We seek to investigate the political slant of the generated responses, i.e. whether a Republican response differs from a Democrat response. Several studies have shown that the particular wording of a text can reveal the author’s political leaning. For instance, what Democrats refer to as “estate tax” is called “death tax” by the Republicans. We hypothesise that, when asked the same question, a Republican would answer differently compared to a Democrat.

The main contributions of this work include:

- We implement a method for creating local knowledge graphs from plain text (congressional speeches).
- We compare the performance of different information retrieval models on these political speeches.
- We investigate how an existing method [8] designed for generating text from knowledge graphs performs in the context of open-domain question answering.
- We analyze the performance of one of the most popular text generation models, namely GPT-2, on the same

task.

- We publish the code for this project on GitHub¹ so that it can be used for further research.

Based on a given question, our method enables a user to retrieve party-dependent speeches, which can be used to create a knowledge graph. The knowledge graph synthesises the information from the speeches by aggregating it and removing duplicate information (duplicate nodes/edges). Furthermore, this knowledge graph can be fed into a language model in order to generate text based on the party-specific information from the graph.

II. RELATED WORK

This paper unites several aspects of NLP, first information retrieval, then creation of knowledge graphs from plain text and lastly text generation from local knowledge graphs.

A. Knowledge Graph Augmentation

The paper from Chen et al.[9] shows how to construct a political opinion-aware knowledge graph based on a pre-existing one. Ideological attributes are incorporated into the existing entities from the original knowledge base. The key difference between their and our approach is that we mine partisan speeches and create separate party-specific knowledge graphs from plain text instead of using an existing knowledge graph. For this reason, the resulting knowledge graphs do not need further induction of partisan slant.

B. Text Generation

A desirable characteristic of a question answering system is the ability to generate answers that are consistent with the question. Yang et al. [6] show that integrating facts from a knowledge base is useful in topic-to-essay generation as it helps generate texts that are more novel, diverse and topic-consistent. They propose a memory-augmented neural model trained using adversarial learning that incorporates knowledge from ConceptNet[10]. The generator is an encoder-decoder model that takes a list of topics as input and produces text consistent with the topics by using an attention mechanism to find the most relevant concepts from the memory matrix. The discriminator is used to evaluate whether the generated text is consistent with the input topics.

The paper of Fan et al. [11] aims to solve the problem of long-range dependency and creativity in story generation. The authors design a convolutional Seq2Seq model with a self-attentive decoder based on a prompt trained from convolutional language model. The generation process uses top-k random sampling for words in order to avoid repetitiveness (which may appear when using beam search) and unlikely words (as in random sampling).

Similarly, Koncel-Kedziorski et al. [8] apply a state-of-the-art information extraction system that is used to

construct knowledge graphs on texts with implicitly strong structural format. In this study, such graph and text features are encoded via a Graph Transformer and a Bi-RNN in the training stage. An attention-based decoder then generates an abstract of scientific papers from encoded embeddings.

C. Evaluation Metrics

Due to the unsupervised nature of text generation, the evaluation stage tends to involve much effort from human experts. Therefore, automatic evaluation metrics are needed in order to thoroughly test the performance of the models.

Diversity and creativity in dialogue response imposes challenges in current unsupervised evaluation metrics. The survey provided by Liu et al. [12] examines pros and cons of two common types of methods in such text generation tasks: word overlap-based (BLEU, ROUGE, METEOR) and word embedding-based (Greedy Matching, Embedding Average and Vector Extrema). Their empirical study shows a low correlation between these metrics and human judgements on certain tasks, and warns against overly using them without carefully considering whether they are suitable for a specific task.

III. DATA AND SUMMARY STATISTICS

The *United States Congressional Records* are archived in a daily edition and a bound edition. The bound edition -from 43rd to 111th (1873-2011)- was chosen as the main corpus for this project since it is a permanent version, obtained from an organised version of the daily edition. These transcripts are assumed to consist of a sufficiently large number of speeches on diverse topics from the two main parties (Republican and Democrat). Meanwhile, these speeches are expected to reflect the divergences in terms of discourse and perspectives of the parties. For example, a Democrat senator uses “estate tax” as in “I am especially disappointed in the provision on the estate tax.”, while a Republican senator addresses the topic with the phrase “death tax” as in “The same people that support death tax seem to be trying to say they inherited every problem that exists.” Such differences are expected to be the key distinguishing feature between the parties, both for generating text that sounds like a Republican/Democrat and for the binary classification of the generated text.

The corpus contains full speeches parsed per speech with metadata such as speakers and their party affiliation. The typewritten text was converted into an electronic format using an optical character recognition (OCR) system, thus we notice non-negligible spelling errors. Particularly, there are frequent errors around hyphens at the end of the original documents and punctuation errors, where full stops got mistaken for commas. These errors have to be corrected in an initial pre-processing stage as they are bound to introduce further errors into the later stages of the pipeline (some

¹<https://github.com/rostrom36/Partisan-Responses>

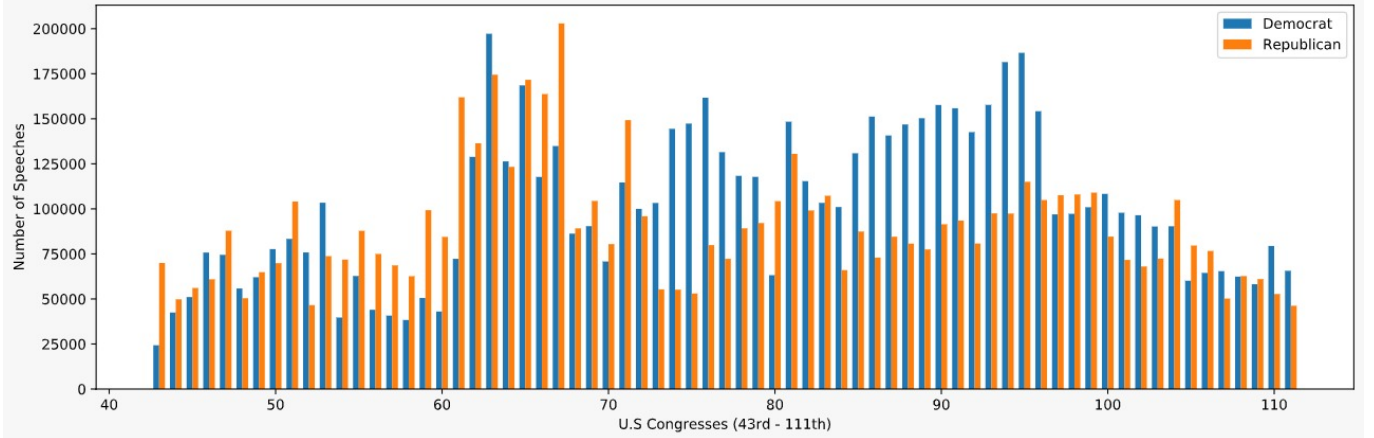


Figure 2. Amount of speeches per Congress. Blue are Democrat speeches, red are Republican speeches.

libraries take whole sentences as input and broken sentences are likely to be processed incorrectly). The dataset contains a total of 13,309,706 speeches from 11 parties. We extract only speeches either from Republican or Democrat speakers, resulting in 13,179,121 remaining samples. The focus on Republican and Democrat speakers only is motivated by their textual abundance and well-known conflicting ideology. The distribution of the number of speeches from the two parties over the timespan 1873-2011 is illustrated by Figure 2. The dataset in total has 46.9% Republican speeches with an average length of 644.37 characters (including spaces and punctuation), and 53.1% Democrat speeches with an average length of 706.21 characters. While the speeches have a reasonable length on average, a large proportion of them are either too short (a single sentence) or too long to be processed by the models and will have to be discarded.

Associated with the corpus, the researchers who provided the dataset also grouped a list of partisan phrases into 22 substantive topics. Each topic contains related keywords and phrases, which were extracted from the congressional speeches in this corpus and some manual additions [13]. These phrases will be used for filtering in order to discard the organisational and other unusable speeches.

IV. METHODS

The tasks of generating party-specific responses for questions can be divided in the following sub-tasks:

- Obtaining a dataset that has the right format for training the language generation model (IV-A)
- Finding the most relevant speeches for a given question (IV-B)
- Creating knowledge graphs using the information extracted from these retrieved speeches (IV-C)
- Generating answers to the questions using the information from the knowledge graphs (IV-D)
- Evaluating the generated responses (IV-E)

For the search stage, the speeches have to be stemmed in order to conflate morphological variants into the same token. This is a time-consuming process, thus the vectorized speeches and their text need to be pre-computed and stored in a database. The most relevant speeches for the questions need to be collected at run time which implies that the knowledge graphs that are based on these speeches are also computed on the fly at run time. Figure 3 gives a visual overview of the sub-tasks outlined in the remainder of this section.

A. Dataset Creation

We ran several experiments, with different datasets constructed starting from the U.S. congressional Record. The last attempt described in this section was used for obtaining the reported results and referred throughout this report.

The first approach was to train a language generation model on paired questions and answers from the U.S. Congressional Records. As many of the speeches are merely organisational (e.g. “Will the Senator yield?”, “Mr. Speaker, I offer a motion to commit.”), they had to be filtered out. In order to obtain this, we used the list of topic-specific phrases associated with the substantive topics identified in Gentzkow et al. [13]. Each speech containing interrogatory sentences from the whole corpus was paired with the next speech, preserving only those that contain at least one phrase from the phrases list. When a speech contains multiple questions it is unclear which one the next speech answers, thus we only preserved speeches containing a single question. We further filtered out pairs whose answers are too short (they contain a single sentence) or too long (they contain more than 50 sentences) and those whose speaker is neither Republican or Democrat.

Due to a lack of sufficient questions, another path that was investigated was creating questions for the speeches in our corpus. To do so, we fine-tuned a T5-Base[15] model

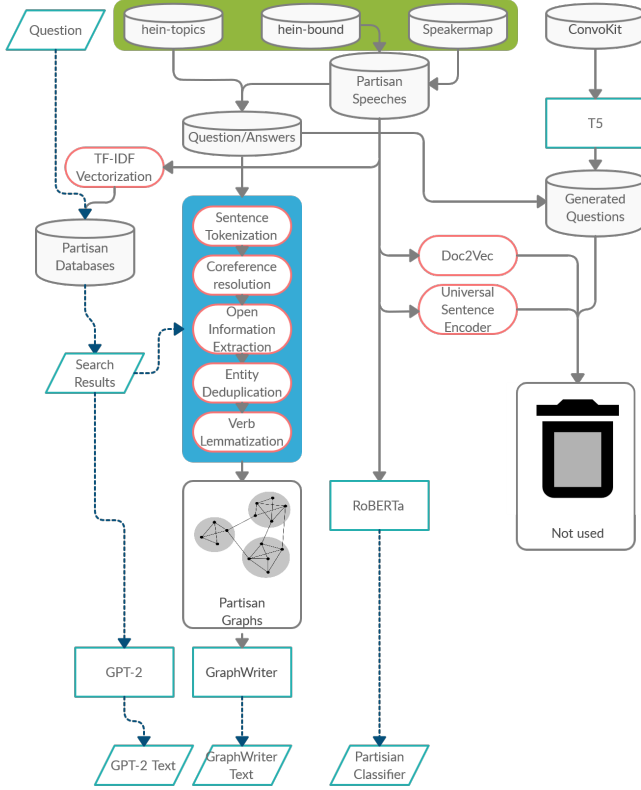


Figure 3. Overview of the processes in the paper for the question-answer dataset. Dotted lines are at run-time, full lines are pre-computed/training. Cylinders are data, small rectangles are models, trapezoids are input/output, ovals are processes. The green data is all bundled in the Stanford data set[14].

using the `simpletransformers`² library. To our knowledge, there is no dataset containing questions and answers from the U.S. Congress, thus a similar one was chosen from Great Britain’s Question Time, which is included in Convokit³[16]. The task was to generate a request (question) for a given answer/speech from a politician. There are obvious differences between the real data (see Table IV) from the congressional debates and the generated questions (see Table III). The generated questions contain titles such as “Honorary Friend”, which are not used in the U.S.. The model cannot avoid learning such titles from the training data, thus this approach was deemed unusable.

From the previous experiments, we found that using only question-answer pairs extracted from U.S. Congressional speeches led to poor results. Thus, we also experimented with maintaining all speeches. Furthermore, the issue with the British titles from the second attempt emphasises the importance of background information about the speeches. In order to keep the size of the dataset manageable for

experiments, we decided to use the speeches from the recent U.S. congressional editions (from the 106th to the 111th). We follow the same approach as in our first attempt for pre-processing. In addition, we remove high-frequency terms such as “Mr. Speaker” and “Madam Speaker” and long organisational sentences like “I yield my time to ...”.

B. Speech Search

The first step consists of extracting the speeches that are the most relevant to a given question. This search step is performed on a filtered subset of the corpus of congressional speeches (see IV-A), following a preprocessing stage. All speeches as well as the questions are stemmed and then vectorized using TF-IDF. For TF-IDF, the most common 1 million unigrams and bigrams are used as features and the rest of the tokens are discarded. We extract the K-Nearest Neighbours of a question by computing the cosine similarity between the question vector and the vectors of the speeches, in an exhaustive search manner. Even though it is reasonably fast, this step could be sped up by using Approximate Nearest Neighbours. We also experiment with this method and it appears to perform well only when the dimensionality of the vector space is not too large (up to 1,000 dimensions) and could be used together with doc2vec or other document embeddings.

C. Knowledge Graph Construction

Most of the well-known knowledge graph databases like ConceptNet[10] contain simple relations like “Pirmin Schwegler –is born in– Ettiswil”. In addition, these databases are too big to be stored locally and thus, it may be tedious to query for a sizeable, dense subgraph. Also, many of the nodes and edges of such a graph are not related to U.S. politics. Adapting any commonly-used knowledge graph (e.g. ConceptNet, DBpedia) to fit the GraphWriter model does not seem to be promising. Therefore, a natural alternative is to create a knowledge graph from scratch using the congressional speeches. On a high level, the goal is to construct a graph by extracting useful triplets $\langle \text{SUBJECT}, \text{OBJECT}^*, \text{PREDICATE}^* \rangle$ from the text and using them as building blocks for the local knowledge graph. The subjects and objects serve as nodes and predicates are directed edges connecting the subject and object. For example, a sentence “The bill will force the taxpayer to pay” will output the triplet $\langle \text{THE BILL}, \text{THE TAXPAYER TO PAY}, \text{FORCE} \rangle$. The detailed local knowledge graph construction is described below. Additionally, the construction requires further text pre-processing due to the software and hardware constraints and the need to remove redundancy.

First, we randomly sample 21,500 speeches from the filtered data (see section IV-A) as this number requires reasonable processing time in experiments. The method is inspired from the work by Fan et al. [17]. A concise graph ideally represents the same references in the same

²<https://github.com/ThilinaRajapakse/simpletransformers>

³<https://convokit.cornell.edu/documentation/parliament.html>

node. Therefore, for each speech, coreference resolution is applied to help collapse co-referred agents into the same graph node (e.g. “he” and the correct agent “the president” together go into node “the president”). We choose the state-of-art coreference implementation from Huggingface, namely `neuralcoref`⁴ because of its reported high performance. Afterwards, in order to reduce sparsity in graph, we use named-entity recognition (NER) to detect entities of the type “person” and “time” and simply mask them with the category name. One motivation would be that the partisan stance is displayed in the discourse regardless of specific senators and time era. For instance, instead of creating a node for “Senator X”, “Senator Y” and “Senator Z” from democratic party respectively, one node “Senator PERSON” will be constructed to connect other parts in the graph. As we seek to extract party-specific opinions, the name of speakers is irrelevant and can be discarded. Furthermore, we argue that this step preserves privacy to some extent. We chain coreference resolution and NER using a `SpaCy` pipeline in order to run NER separately on the results of coreference resolution. Finally, `AllenNLP`’s implementation of open information extraction⁵ is used to parse the sentences and generate `<SUBJECT, OBJECT*, PREDICATE*>` triplets, which should be an extract of the information provided in the sentence. In order to avoid inconsistencies, triplets which have more than one object, more than one predicate or auxiliary verbs with the exception of “to be” are discarded.

These triplets are the building blocks for the knowledge graph which are created using the graph library `NetworkX`⁶. To make the graph denser, the predicates are lemmatized using `spaCy`’s lemmatizer⁷ and the entities are deduplicated by computing their similarity using `Universal Sentence Encoder`⁸. If the similarity between a new node and an existing node is less than a threshold set empirically, we do not create a new node. To keep track of the already existing nodes, we use `hnswlib`⁹ as it best allows to build a database while still using it.

D. Response Generation

1) *GraphWriter*: The workhorse for this task is *GraphWriter*¹⁰, which was originally trained for generating abstracts for scientific articles. In the original form, *GraphWriter* uses a processed version of the *AGENDA* dataset which has 7 possible relations and 5 types of entities. By contrast, our data contains more than 1000 relations (verbs) and only one generic type of entity (“phrase”), resulting in a much more complicated graph structure for training. At

generation time, *GraphWriter* uses its graph encoder-decoder network with four attention heads to obtain probabilities, which are used in a beam search fashion to produce the text.

As there does not exist a pre-processing script for the data in the *GraphWriter* repository and the paper does not offer details on it, the pre-processing stage was reverse-engineered. The pre-processing step was implemented as described in the section IV-C. To reduce the resource intensity, coreference resolution was run ahead of open information extraction (intermediate results were stored) but more care was needed for constructing the knowledge graphs as *GraphWriter* does not use nested relations while training. Using this additional criteria, we perform another filtering step on the speeches from the dataset curated in subsection IV-A, making sure we also drop those that are too long or do not contain any relations. The final dataset contains around 10,000 speeches and we use a train-test split similar to the one from the original paper, with 1’000 validation samples and the rest as training samples.

We train *GraphWriter* for 40 epochs with a learning rate of 0.1, then the checkpoint with the lowest validation loss is chosen and trained for another 20 epochs with a learning rate of 0.05. For generation we choose the model with the lowest validation loss.

2) *GPT-2*: We compare the performance of *GraphWriter* to one of the most powerful models for text generation, namely *GPT-2* [21]. We fine-tune the pre-trained model on the datasets filtered by partisan labels from section IV-A, one for Republicans and one for Democrats, in order to compare the responses generated for the same question and check whether it manages to learn any party-specific information. Care must be taken so that *GPT-2* is trained for long enough in order to indeed capture political ideologies without overfitting. We notice from experiments that overfitting would result in big chunks of text being naively copied from the training data. Due to the limited RAM available in Google Colab, it was trained on approximately 100,000 samples, for 2,000 steps.

E. Response Evaluation

Due to the creativity nature of our task, ideal evaluation should be conducted thoroughly by human evaluators, preferably domain experts in political science. Within our capabilities, we designed the evaluation metric to consist of (1) manual evaluation (on our own), and (2) two machine classifiers for distinguish parties from the generated text. Details regarding both classifiers are provided below.

To evaluate the political slant of responses, a *RoBERTa*-base model [22] was fine-tuned on the speeches from IV-A with the help of the `simpletransformers` library¹¹ (with hyperparameters `max_seq_length=128`, `sliding_window`, `learning_rate=4e-6`). *RoBERTa* is an adaption on the widely-used

⁴<https://github.com/huggingface/neuralcoref>

⁵[https://demo.allennlp.org/open-information-extraction\[18\]](https://demo.allennlp.org/open-information-extraction[18])

⁶<https://networkx.github.io/>

⁷<https://spacy.io/api/lemmatizer>

⁸[https://tfhub.dev/google/universal-sentence-encoder/4\[19\]](https://tfhub.dev/google/universal-sentence-encoder/4[19])

⁹[https://github.com/nmslib/hnswlib\[20\]](https://github.com/nmslib/hnswlib[20])

¹⁰[https://github.com/rikdz/GraphWriter/\[8\]](https://github.com/rikdz/GraphWriter/[8])

¹¹<https://github.com/ThilinaRajapakse/simpletransformers>

Transformer model BERT[2] which is expected to perform better as it was trained for longer, with additional data (news, web text and a story corpus), using slightly changed training objectives, and an adapted masking technique. On our dataset, the best accuracy achieved was 71%, which could potentially be improved further by using more data and experimenting with more combinations of hyperparameters. However, the poor quality of the responses generated using GraphWriter did not allow for partisanship classification, therefore we did not proceed with further improvements of this RoBERTa model.

In addition to RoBERTa, we implemented a lexicon-based method. The U.S. Congressional Record corpus provides a list of the 1,000 most partisan phrases for each session of the Congress, together with scores for partisanship. The average of these scores was used to compute an overall partisanship score for each speech. The accuracy using this method proved to be only slightly higher than random guess, thus we did not proceed with it further. On a closer inspection, it could be noticed that the partisanship scores for the phrases, as well as their scores vary widely between each year of the congress and a large proportion of them cannot be assigned consistently to one party, thus being of little use for classification.

V. RESULTS

A. Dataset

Appendix C displays some pairs of questions and answers from the dataset created using the U.S. Congressional Record. In general, the attempted question-answer method described in Section IV-A manages to filter out organisational and rhetorical questions. Moreover the speeches answer their paired questions. However, these pairs lack the advantages of a dataset designed for question answering, where the answers are generally short and concise paragraphs. The parliamentary speeches often contain formalities such as “Mr. President, I, again, will be happy to attempt to clarify this issue for my colleague and friend.” and they are rather long, containing an answer to the question but also additional unrelated information.

B. Speech Search

Among TF-IDF, doc2vec, Universal Sentence Encoder, the most accurate search results were obtained when using TF-IDF. All the three methods were tested on the same questions and the search results were evaluated using manual inspection. The search results returned when using doc2vec are all short answers that are unrelated to the question. The speeches returned when using USE are on the same topic as the question but often do not contain an answer to it. The reason for this might be the fact that the corpus contains a large proportion of very long speeches, which are hard to transform in a low-dimensional representation. Examples of search results for the three methods are shown

in Appendix B. Table IX shows some examples of questions, together with one of the retrieved speeches for each party using TF-IDF. The list of questions on which the search was tested on was created manually starting from issues that Republicans and Democrats generally disagree on. For testing purposes, we also added some questions related to topics on which both parties presumably share similar views. There are indeed noticeable differences between the results for the two parties and in most cases the speakers appear to have opposing views that match their parties’ stance.

C. Knowledge Graphs

Figure 4 shows a knowledge graph for one speech that answers the question “Should abortion be illegal?”. Due to the complicated nature of the discourse specific to the congressional debates, the nodes/entities obtained using open information extraction are often long and non-trivial. For this reason, the language generation model might have trouble when producing text using these nodes. This also adds to the sparsity of the graph, which is apparent by looking at the many unconnected shards. In a best-case scenario, the graph would be quite dense, so that a text generator has many meaningful nodes neighbouring an entity to choose from. One remedy for this problem would be using more speeches for creating a single knowledge graph. However, this would result in a bigger graph, which would be harder to handle using the limited resources from Google Colab (sometimes even when using a single speech we run into memory issues).

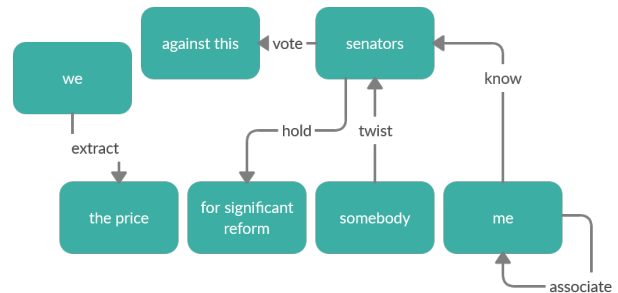


Figure 4. Example of knowledge graph constructed from one of the speeches. Entities (nodes) are connected with each other with directed edges, which represent verbs. The edges are named according to the stemmed verb.

D. Response Generation

Two methods were used for text generation, namely GPT-2 and GraphWriter, both discussed in the following subsections.

1) *GPT-2*: In general, the GPT-2 model fine-tuned on the U.S. Congressional Record produces coherent and grammatically correct text, but it sometimes contains repetitive

sentences. Even though it is a very good model for language generation, it can be easily noticed that using GPT-2 for open-domain question answering is far from an ideal choice. As it was not originally designed for this task, its architecture only allows to generate text after a prompt (the question). Even though through fine-tuning it manages to learn some political slant, the responses do not directly address the question. The generated text is generally on the same topic as the input question but the model lacks the knowledge required for specifically answering the questions.

The generated responses, which can be seen in Figure I, often contain stereotypical opinions of the 2 parties (Republicans/Democrats) but sometimes the answers can also be neutral or display the views of the opposing party. This is mostly due to the training data, which contains a large proportion of speeches that do not display any political ideology and speeches that contain citations (often from people of the opposing party). For this reason, the model is unable to distinguish between the cited speech and the main speech and thus it might capture conflicting views.

2) *GraphWriter*: Even though using information from knowledge graphs has important advantages when performing text generation, the limitations of the chosen graph model seem to outweigh the benefits. As it can be noticed in Table II, GraphWriter fails to generate coherent responses to the questions. A more extensive table of examples can be found in Appendix section E. The generated responses have a lot of repetitions and in most cases they are barely coherent. While the performance is undeniably poor, the current responses start to display support or objection on the issues present in the questions.

Despite carefully pre-processing the data in several stages, the final training set for GraphWriter still has several shortcomings which are very likely to have negatively impacted the training of the model. Firstly, the first example shows the limitations of our sentence removal such that “Mr. Chairman,” was not removed due to the inserted space from OCR. Furthermore, organization entities such as “department of defense” or some committees increase the graph sparsity by occurring frequently, yet connecting to few other graph nodes. Another important entity, “PERSON”, does not include phrases like “the gentleman from New York” which results into too many unnecessary graph nodes. We believe that better further lemmatization may address these issues.

Besides the training data, GraphWriter itself needs additional future work to adjust its architecture to be more suitable for open-domain question answering. Originally, it was trained on much simpler, more semantically structured, and less rhetorical text than the speeches from our dataset (U.S. Congressional Record). The current implementation of GraphWriter makes it hard to fit the samples into the VRAM available in Google Colab. For this reason, many useful speeches that are beyond a certain length get thrown away. The remaining speeches are shorter (at most 1,000 charac-

Question:	Should we introduce more gun control measures?
Republican Answer:	No, I think it would be really wrong to do so. That is why I want to go back to the issue of the price of guns. If you look at the background of the gun control legislation, the background that we passed in 1996 with regard to background checks was that we were trying to provide people with a gun. We do not want to hear anybody say that this is the right way to do it. We want to provide some sort of background check for people who have a gun, because that is what we do, and so we want to make sure that those who are at risk of being shot by a gun are not going to be able to purchase a firearm. We do not want to hear anybody saying that. It is not the right way to do it. We want to make sure that those who are at risk of being shot by a gun are not at risk of being able to purchase a firearm.
Democrat Answer:	There is a chance that we could be in the position of having a second term in this country when we are considering the next administration, I urge our colleagues to support the amendment we have offered and engage in a dialogue with the American people. The American people want us to do that.
Question:	What is the biggest threat to America?
Republican Answer:	We should not make that kind of a mistake. In fact, there is another area of concern. In the area that is most under threat, we have our own military. We have our own defense. We have our own resources. We have our own energy. We have our own technology. We have to protect ourselves. In this segment of the country, we have a lot of other things. We have a lot of other things besides weapons of mass destruction. We have a lot of nuclear technologies. We have a lot of other things, too.
Democrat Answer:	The biggest threat to the United States is the proliferation of a nuclear arsenal. All of us know that if we do not take the steps necessary to protect our citizens, we will be exposed to nuclear weapons. But we also know that there are other nuclear threats that are more fundamental to this country and that are more likely to cause us to start off a nuclear war. The threat of nuclear weapons is a very real and real threat to our national security. Our national defense is not just about our national security. It is about our national security and the security of our country. More than anything else, we are a nation of security. We are a nation of liberty. We are a nation of opportunity, I believe that we have a very serious national security threat, I do not think that we can afford to wait for this Congress, and we have to act quickly, to provide the resources that we need to do that.

Table I
EXAMPLES OF QUESTIONS, TOGETHER WITH GENERATED ANSWERS USING GPT-2, FINE-TUNED SEPARATELY ON REPUBLICAN AND DEMOCRAT DATA. REPETITIVE SENTENCES WERE CUT.

ters) and often contain uninformative sentences. Moreover, the model is incapable of coping with the large number of possible verbs/relations, which cannot be avoided in open-domain question answering. Therefore, with the current architecture, GraphWriter produces low-quality text. A more robust model would be required for a system that aims to answer any question.

- [6] P. Yang, L. Li, F. Luo, T. Liu, and X. Sun, "Enhancing topic-to-essay generation with external commonsense knowledge," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics, Jul. 2019, pp. 2002–2012. [Online]. Available: <https://www.aclweb.org/anthology/P19-1193>
- [7] H. Hayashi, Z. Hu, C. Xiong, and G. Neubig, "Latent relation language models," 2019.
- [8] R. Koncel-Kedziorski, D. Bekal, Y. Luan, M. Lapata, and H. Hajishirzi, "Text generation from knowledge graphs with graph transformers," 2019.
- [9] W. Chen, X. Zhang, T. Wang, B. Yang, and Y. Li, "Opinion-aware knowledge graph for political ideology detection," in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, 2017, pp. 3647–3653. [Online]. Available: <https://doi.org/10.24963/ijcai.2017/510>
- [10] R. Speer, J. Chin, and C. Havasi, "Conceptnet 5.5: An open multilingual graph of general knowledge," pp. 4444–4451, 2017. [Online]. Available: <http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14972>
- [11] A. Fan, M. Lewis, and Y. Dauphin, "Hierarchical neural story generation," 2018.
- [12] C.-W. Liu, R. Lowe, I. V. Serban, M. Noseworthy, L. Charlin, and J. Pineau, "How not to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation," 2016.
- [13] M. Gentzkow, J. M. Shapiro, and M. Taddy, "Measuring group differences in high-dimensional choices: Method and application to congressional speech," *Econometrica*, vol. 87, no. 4, pp. 1307–1340, 2019. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA16566>
- [14] M. T. Matthew Gentzkow, Jesse M Shapiro, "Congressional record for the 43rd-114th congresses: Parsed speeches and phrase counts," 2018. [Online]. Available: https://data.stanford.edu/congress_text
- [15] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," *arXiv e-prints*, 2019.
- [16] C. D.-N.-M. Justine Zhang, Arthur Spirling, "Asking too much? the rhetorical role of questions in political discourse," in *Proceedings of EMNLP*, 2017.
- [17] A. Fan, C. Gardent, C. Braud, and A. Bordes, "Using local knowledge graph construction to scale seq2seq models to multi-document inputs," 2019.
- [18] G. Stanovsky, J. Michael, L. Zettlemoyer, and I. Dagan, "Supervised open information extraction," pp. 885–895, 01 2018.
- [19] D. Cer, Y. Yang, S. Kong, N. Hua, N. Limtiaco, R. S. John, N. Constant, M. Guajardo-Cespedes, S. Yuan, C. Tar, Y. Sung, B. Strope, and R. Kurzweil, "Universal sentence encoder," *CoRR*, vol. abs/1803.11175, 2018. [Online]. Available: <http://arxiv.org/abs/1803.11175>
- [20] Y. A. Malkov and D. A. Yashunin, "Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs," *CoRR*, vol. abs/1603.09320, 2016. [Online]. Available: <http://arxiv.org/abs/1603.09320>
- [21] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language models are unsupervised multitask learners," 2019.
- [22] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "Roberta: A robustly optimized BERT pretraining approach," *CoRR*, vol. abs/1907.11692, 2019. [Online]. Available: <http://arxiv.org/abs/1907.11692>

APPENDIX A. GENERATED QUESTIONS

The first part of the generated questions from Section IV-A cover a test made with British context (Table III) and the original request. In Table IV is a context and a request generated from the American corpus.

APPENDIX B. SEARCH RESULTS USING DIFFERENT METHODS

We show how the search step performs using three different methods : TF-IDF, doc2vec¹² and Universal Sentence Encoder¹³. Examples of questions together with the retrieved speeches using the three methods are shown in Table V, Table VI and Table VII.

APPENDIX C. DATASET

Table VIII shows some examples of question and answer pairs obtained using the procedure described in Section IV.

APPENDIX D. SPEECH SEARCH

Table IX shows found speeches by searching for the question using TF-IDF as it is discussed in Section V-B, there is one example for both parties per question.

APPENDIX E. GENERATED RESPONSES

Table X shows some more GraphWriter responses. The discussion to it can be found in Section V-B.

¹²<https://radimrehurek.com/gensim/models/doc2vec.html>

¹³<https://tfhub.dev/google/universal-sentence-encoder/4>[19]

Answer/ Context	The special fund of £ 34 million that we have provided for the historical inquiries team to look into the unresolved murders of the years of the troubles is , of course , additional money that does not come out of the day - to - day policing budget . It is important that we focus all the resources that we can on day - to - day policing—that is , on neighbourhood policing and dealing with antisocial behaviour , for example—but we must also deal with the remaining threats from dissident republicans , loyalist paramilitaries or those involved in organised crime networks . I know that the hon Gentleman cares very deeply about such matters , and that he will agree that we must remain very strongly focused on them .
Question/ Gener- ated:	Although the hon Gentleman knows that his Government are on the right side of their budgetary policy—north of 50p—will he comment on the general welfare measures that the Home Secretary and Cabinet have proposed to fund?
Question/ Truth:	Does not the Minister 's very proper condemnation of the events of yesterday and the weekend underline the necessity for keeping an absolutely first - class police force in Northern Ireland , under the inspired leadership of an admirable Chief Constable ? Will he ring - fence expenditure on the past and try to ensure that is totally separate from the running budget for police needs ?

Table III
EXAMPLE OF GENERATED QUESTIONS IN COMPARISON TO THE TRUTH GIVEN AN ANSWER FROM THE BRITISH DATA.

Answer/ Context	Mr. Speaker, I thank the gentleman. Mr. Speaker, you dont have to look far to see that families across this country are gripped with a tremendous amount of fear and uncertainty. They fear for their jobs, if they have one. They fear for their future as they see their 401(k)s, their college savings accounts collapse. They fear that their elected leaders dont get it. They fear that this Congress may very well be incapable of change, incapable of producing the kind of results that they want and to get it right. Under existing House rules, when a bill is brought to the floor that includes a tax increase, the minority has a right to offer a motion to strike that increase, and the Republican minority had done that on nearly half a dozen occasions over the past 2 years. With this rule change now, though. House Democrats are trying to push through what we Republicans will no longer have, the ability to say “no” to higher taxes. We will not be able to simply strike a tax increase and demand an up or down vote. In fact, the only option we will have would be to replace one tax increase with another. There will be no ability for us to cut taxes to lighten the burden on the middleclass families that are hurting right now. One can see that this rule change makes it a lot easier for the Democrat majority to In fact hide tax increases inside other larger bills. In fact, that is why all of us are sitting here scratching our heads. If the House Democrats feel a tax increase is necessary, then why wouldnt they allow for a full and open debate? Why not let the American people have a say? Why not let the hardworking people of this country hear why Washington is once again looking to take more of their hard earned money? Either way, what is clear, this type of partisan rules change flies in the face of a new era of openness and transparency that Presidentelect Obama has promised, I take the Presidentelect at his word, I believe he wants transparency, openness, and debate, I believe he wants Washington to begin to do business differently, I believe he is serious in wanting Congress to work together for the good of all of our constituents. But apparently that word hasnt made its way down to the leadership of the House.
Question/ Gener- ated:	Will the Prime Minister explain to the House the plight of thousands of students and parents as a result of economic recession and, if possible, to some employers, teachers, doctors and other social workers who may have to live with family

Table IV
EXAMPLE OF GENERATED QUESTION GIVEN AN ANSWER FROM THE AMERICAN DATA.

Question:	What reforms were adopted by the 110th Congress?
Speech:	... “ The Omnibus Public Land Management Act of 2009 is a collection of over 150 public land bills that were reported from the Committee on Energy and Natural Resources during the 110th Congress, for which we have not been able to get unanimous consent to take up and pass during the 110th Congress, I have included them in the Omnibus Public Land Management Act of 2009 to facilitate their early consideration in the new Congress, and not “primarily at the request of a Senator.” Nevertheless, even though no Senator has specifically requested me to include a congressionally directed spending item in the Omnibus Public Land Management Act of 2009, in the interest of furthering the transparency and accountability of the legislative process, I have posted on the Web site of the Committee on Energy and Natural Resources a complete list of all provisions in the Omnibus Public Land Management Act of 2009 that authorize a specific amount of spending authority that is targeted to a specific State or locality, other than through a statutory or administrative formuladriven or competitive award process. The list includes the name of the principal sponsors of the Senate bills in the 110th Congress that have been incorporated in the Omnibus Public Land Management Act. In addition, I have added several other nonpublicland measures from the 110th Congress at the request of the majority leader. Most of these provisions were included in the Advancing Americas Priorities ActS, 3297in the 110th Congress. They include: the Christopher and Dana Reeve Paralysis Act, subtitle B of title I of S, 3297, four parts of subtitle B, relating to oceans, of title V of S, 3297, and title VII of S, 3297, relating to the authorization of a greenhouse facility for the Smithsonian Institution. These provisions were determined not to constitute “congressionally directed spending items” in the Advancing Amercias Priorities Act. See 154 Cong. Rec, 1657374. July 26, 2008. In addition, I have added the Coastal and Estuarine Land Conservation Program Act. H.R, 1907 in the 110th Congress, and the Smithsonian Institution Facilities Authorization Act of 2008. H.R, 6627 in the 110th Congress, at the request of the majority leader. The grant program established under Coastal and Estuarine Land Conservation Program Act, section 12507 in the Omnibus Public Land Management Act, does not constitute a congressionally directed spending item because the funds are to be allocated through a competitive grant process. The authorizations in the Smithsonian Institution Facilities Authorization Act, sections 15101 and 15102 of the Omnibus Public Land Management Act, do not appear to constitute congressionally directed spending items because they were requested by the Board of Regents of the Smithsonian Institution, and because they originated in the House of Representatives, where the committees of jurisdiction determined they did not constitute congressional earmarks. See H. Rept, 110842, part 1, at 5, 2008. Committee on House Administration, and H. Rept, 110282, part 2, at 4, 2008. Committee on Transportation and Infrastructure. Finally, I have added the ShoshonePaiute Tribes of the Duck Valley Ieservation Water Rights Settlement Act. H.R, 5293 in the 110th Congress, at the request of the majority leader. This act ratifies a water rights settlement among the ShoshonePaiute Tribes of the Duck Valley Reservation, individual water users, and the State of Nevada. Section 8 of H.R, 5293, section 10807 of the Omnibus Public Land Management Act, creates two trust funds to settle the legal claims of the ShoshonePaiute Tribes against the United States for compromising tribal water rights and failing to maintain the Duck Valley Indian Irrigation Project. ...
Speech:	Madam Speaker, let me congratulate you for your reelection as Speaker of the House. It is an honor that you tave served with great distinction and verve, I look forward to more of your continued leadership in the 111th Congress. Mr. Speaker, I rise today in support of H. Res, 5. Adopting the rules for the One Hundred Eleventh Congress. The House Rules Package provides commonsense reforms that will enable Congress to work more efficiently for America. In the 110th Congress. Democrats put forth critical measures to restore integrity and accountability to the House. These reforms were the most sweeping ethics and lobbying reforms since Watergate and has changed the way Congress does business in Washington. The reforms adopted by the 110th Congress included banning gifts from lobbyists, prohibiting the use of corporate jets, mandating ethics training for all House employees, establishing a new, independent Office of Congressional Ethics, and ensuring transparency for budget earmarks by requiring the full disclosure of earmarks in all bills and conference reports. The Rule Package for the 111th Congress builds upon these reforms to further strengthen the integrity of Congress. Key provisions include closing the loophole regarding “lameduck” Members negotiating postCongressional employment, codifying additional earmark reforms adopted in midterm in the 1 10th Congress, continuing the Office of Congressional Ethics, maintaining strong PAYGO rules, and improving Congress effectiveness by removing an abusive practice where popular measures are killed through unrelated. “gotcha” amendments on motions to recommit.
Question:	What do you believe about tax increases?
Speech:	Madam Speaker, today, the House is again going to consider the Democrat SCHIP Expansion Act. This polarizing expansion of the welfare state was stopped by the veto pen last year, and thats where It ought to stay. Once again, the bill fails to ensure that all lowincome children will be covered, but, instead, it is going to increase the number of adults on SCHIP. Once again, the bill will grant health care benefits to illegal aliens. The bill will force the taxpayer to pay the health care premiums for children who already have quality, private health insurance. Even though Presidentelect Obama promised we wouldnt tax anyone making less than \$250.000 a year, this bill will impose the most regressive tax on the poor in history with the tobacco tax increase. It is shameful that the Democrats are playing politics with Americas needy children, I urge all of my colleagues to vote against this bill.
Speech:	Madam Speaker, I think that, I would hope, that not only my colleagues but the American people realize that this bill today contains a \$72 billion tax increase on the American people, what Congressional Research Service calls the most regressive of taxes, because it is tobacco taxes. But this is a tax increase that is coming full steam ahead at us. And. Madam Speaker, it is not there to go into a program that we all originally supported the way SCHIP was originally set up. This expanded SCHIP goes to middleincome children, it does not focus on low income and uninsured children. That is a sad day for us. Indeed, part of the 900.000 children that are expected to be added already have access to health insurance, I would encourage all of my colleagues to vote against the tax increase and vote “no.”

Table V
EXAMPLES OF QUESTIONS, TOGETHER WITH TWO OF THE NEIGHBOURING SPEECHES
USING TF-IDF. THE HIGHLIGHTED TEXT SHOWS THE ANSWER IDENTIFIED USING MANUAL INSPECTION.

Question:	What reforms were adopted by the 110th Congress?
Speech:	The majority leader is recognized.
Speech:	The Senator from South Carolina.
Question:	What do you believe about tax increases?
Speech:	Mr. President, the next vote will be on Monday at about 5:30 p.m.
Speech:	Mr. President, I anticipate about 5 minutes.

Table VI
EXAMPLES OF QUESTIONS, TOGETHER WITH 2 OF THE RETRIEVED SPEECHES USING DOC2VEC.

Question:	What reforms were adopted by the 110th Congress?
Speech:	Madam Speaker, let me congratulate you for your reelection as Speaker of the House. It is an honor that you have served with great distinction and verve, I look forward to more of your continued leadership in the 111th Congress. Mr. Speaker, I rise today in support of H. Res. 5. Adopting the rules for the One Hundred Eleventh Congress. The House Rules Package provides commonsense reforms that will enable Congress to work more efficiently for America. In the 110th Congress, Democrats put forth critical measures to restore integrity and accountability to the House. These reforms were the most sweeping ethics and lobbying reforms since Watergate and has changed the way Congress does business in Washington. The reforms adopted by the 110th Congress included banning gifts from lobbyists, prohibiting the use of corporate jets, mandating ethics training for all House employees, establishing a new, independent Office of Congressional Ethics, and ensuring transparency for budget earmarks by requiring the full disclosure of earmarks in all bills and conference reports. The Rule Package for the 111th Congress builds upon these reforms to further strengthen the integrity of Congress. Key provisions include closing the loophole regarding "lame duck" Members negotiating post-Congressional employment, codifying additional earmark reforms adopted in midterm in the 110th Congress, continuing the Office of Congressional Ethics, maintaining strong PAYGO rules, and improving Congress's effectiveness by removing an abusive practice where popular measures are killed through unrelated, "gotcha" amendments on motions to recommit. On this last point, noted Congressional scholar Norm Ornstein pointed out in the Roll Call. August 13, 2007. "Using promptly . . . is a subterfuge, a way to kill bills, and reflects a desire not to legislate but embarrass vulnerable majority Members through a "gotcha" process. The Rules Package protects the minority and still preserves its ability to recommit. Specifically, the minority can offer a motion to recommit "forthwith." where the GOP amendment is immediately voted upon and, if adopted, is added to the bill. Additionally, the minority can offer a straight motion to recommit the bill to committee (in which case the vote occurs on the merits of the bill itself). Mr. Speaker, the Rules Package removes term limits for Committee Chairmen from House Rules. Instead, each party should determine its own rules on the tenure of Committee Chairs and/or Ranking Members and they should be reflected in Democratic Caucus Rules and Republican Conference Rules. In practice, term limits have resulted in the creation of a "payto play" system, where the chief criterion for being selected as a new Chair has in many instances been a Member's fundraising prowess. This had the effect of focusing upon fundraising and undermining the integrity of Congress and the legislative process. Lastly, I am pleased that the Select Committee on Energy Independence and Global Warming, the Tom Lantos Human Rights Commission, and the House Democracy Assistance Commission will be continued. These entities have done tremendous work, I urge my colleagues to support the Rules Package, I believe this package restores integrity and accountability.
Speech:	Under the previous order, three-fifths of the Senators duly chosen and sworn having voted in the affirmative, the bill is passed. The bill was passed, as follows: S. 181 Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled.
Question:	What do you believe about tax increases?
Speech:	Thank you. Congressman AKIN. And you know, when you showed the chart over there about the revenue, the dip and then ultimately as taxes were cut. Federal revenues actually increased. The same thing happened under President Reagan when President Reagan cut taxes, I think one of the myths that is out there is that the deficit grew. Some people tried to attribute that to the tax cuts. But if you really go and look, you'll see a similar chart, you'll actually see an increase in revenue. Unfortunately, you had a Democratic-controlled Congress that spent even more money than the new money that did come in. But in fact, more money came in as taxes were cut. And so I hope we use history as a guide. As you talked about earlier, there is no bill filed yet on this economic stimulus plan. We are expecting in the next week to possibly 2 weeks, there will be a bill filed. And unfortunately, right now what you've got is a bidding war. What started off as maybe a \$400 or \$500 billion proposal has now reached over \$1 trillion where the proposals that were hearing now are \$1.3 trillion.
Speech:	Madam President, what this bill does is probably one of the most important things we need to do in this economic debate, and it is stop the planned tax increases that are going to happen in 2011 for every American. The large score that is being thrown around here assumes we are going to let those taxes go up, but we are not. This is a misrepresentation of the cost of this bill. This bill stops the current tax increases that are planned in 2011, keeps the current tax rate the same. The only change it makes is it lowers the top marginal rate from 35 to 25 percent for businesses, for investors, and for individual Americans. We call it the American option because it leaves money in the hands of the American people and businesses, rather than bringing it to Washington and distributing it our way, I encourage everyone to stop the planned tax increases with the American option.

Table VII
EXAMPLES OF QUESTIONS, TOGETHER WITH TWO OF THE RETRIEVED SPEECHES USING UNIVERSAL SENTENCE ENCODER.

Question:	Do you think that any increases in military spending that we have made should have been cut as well?
Answer:	If I can reclaim my time. Mr. Speaker, of course, I think that we have yet to deal, by the way, with the Department of Defense Appropriations bill, I know that it is going to be marked up, I anxiously look forward to seeing what this will consist of. But frankly, in the proposed budget I believe that that has, as an increase, one of the smallest levels of increases compared to the 22 percent increase that we saw on other appropriations bills. The fact is there is a role for the Federal Government. The number one priority of the Federal Government happens to be the national security of the United States of America. And so to say that because we might have an increase in the level of defense expenditures, as we live in a very dangerous world, and that somehow justifies a multibillion dollar increase in deficit spending is apples and oranges when one looks at what should happen. So I would like to engage, if I might, with my friend from Laredo and say, as I look forward to yielding to him, that as we look at this challenge that families face when they are at the kitchen table, recognizing that with the difficult economic times that we have, they have to rein in their spending. They have to pay down their debts. They have to increase their level of savings. How is it that we, in this measure, can exempt 40 percent the discretionary spending level that is here? How is it that we can say that reducing rates on things like capital gains, which doubled the flow of revenues to the Federal Treasury when reduced by 5 percent, how is it that we can't do that any longer under this so-called PAYGO provision? And I would be happy to yield to my friend if he would like to respond.
Question:	So going back to my question, if you cannot control costs, can you sustain a health care system?
Answer:	In my opinion, no, because, again, if you can't do it for a smaller system, how can you enlarge the system and somehow make it mysteriously work, particularly when there are no models? Massachusetts. Tennessee. TennCare, and so on and so forth, no one has an example of a government-run system that works.
Question:	With respect to the provisions granting the states authority to take enforcement action, is it your intent the states limit their enforcement actions under the new mortgage standards promulgated by the FTC, or under TILA, only to those mortgage industry participants that are not currently supervised by the federal banking agencies or are not Federal credit unions?
Answer:	Yes, the Senator from Connecticut is correct. Our intention was to permit state attorneys general to bring civil actions only against mortgage industry participants that are not supervised by the Federal banking agencies or are not Federal credit unions.
Question:	Congressman from the great State of Missouri, how do you understand this critical issue of D.C. voting rights?
Answer:	Thank you. Congressman ELLISON. One of the most significant measures to find its way into the United States Congress is legislation put forth by our colleague. ELEANOR HOLMES NORTON, who is the delegate for the District of Columbia. This legislation would allow the citizens of the United States of America, who live in the District of Columbia, to finally, to finally, after more than 200 years, have the opportunity to cast their vote to place a representative in the United States Congress. This is a city of almost 600,000 people, and many people around the Nation may be surprised to learn that the District of Columbia is the only city in the United States that must submit its municipal budget to the United States Congress. That, in and of itself, is an injustice. That means that this city, unlike any other city, is subservient to the Congress of the United States and they have no voice whatsoever. The sad thing goes further. Forty percent of the District of Columbia own their own homes, and coming from those homes are young men and women who have died in the world wars, who have died in Vietnam and who are still dying in Iraq and Afghanistan.
Question:	Mr. President, if I could further inquire of my colleague and I thank him for his response has not the Speaker of the House indicated there is no intention of including a reconciliation instruction with respect to climate change in the House provisions?
Answer:	Might I further inquire. Mr. President, of my colleague: Wouldn't it be logical for me to object if my conclusion is that there is not going to be any reconciliation instruction with respect to cap and trade? We don't have it in the Senate resolution before us. The Speaker of the House has made clear they are not going to have a reconciliation instruction to be used in the House with respect to climate change legislation, I must say, I understand the concern, but I don't think there is a basis for it, I don't think there is a prospect that there is going to be the use of reconciliation for the purpose of climate change resolution coming back from the conference committee. It is not in the Senate, the Speaker has made clear they do not intend to use it on the House side, so I would just say to my colleagues that I understand the concern, I understand making the point that is fully legitimate but I don't think it is a concern that is based on what is going to happen. There are plans on the House side to use reconciliation for health care and for education. That clearly is part of their resolution. Not part of ours, but part of theirs. So I have to say to my colleagues, I don't think there is a basis for concern that the reconciliation process is going to be used for climate change legislation.

Table VIII
EXAMPLES OF QUESTIONS AND ANSWERS EXTRACTED FROM THE U.S. CONGRESSIONAL RECORD.

Question:	Should abortion be illegal?
Republican result:	<p>Mr. President, I urge all my colleagues to read Debbie's letter describing her experience with two abortions. Abortion is one of the, if not the, most serious violations of human rights that one person can inflict on another. It is the deliberate destruction of the offspring of a human being which is without a doubt a human being. But to literally add insult to injury, abortionists are degrading women and endangering the health of the mothers by not adequately informing them of the risks and nature of abortion. Pregnant women, 4,100 of them each day, are procuring abortions, and most of them will not know the basic facts regarding the gestation of their infant, or the type of abortion procedure to be used. It is imperative that S. 2791, my bill requiring informed consent before abortion, be enacted into law in order to protect the emotional and physical health of this Nation's women. I ask, in particular, my colleagues from Massachusetts to heed the plea of Debbie, their constituent, and support informed consent before abortion. I ask that Debbie's letter be printed in the RECORD. The letter follows: DEAR SENATOR HUMPHREY: My first abortion was illegal at the age of 18. That didn't stop the doctor from asking me if I wanted an abortion. That was all the counseling I received. It is tragic that the counseling most women considering abortion receive is so nearsighted. Those who counsel women at abortion clinics are simply interested in rushing scared women and girls into abortions which many will regret later. It may take months, even years, but once the realization of what abortion is and does sets in, women suffer devastating emotional crises and find it very hard to forgive themselves for the role which they had in killing their own children. The tragedy of killing an unborn child by abortion is compounded by the physical harm done by abortion to many women who undergo the procedure. I myself suffered from a very bad infection after my abortion, an infection which spread throughout my body. Other women I know have been forced to have total hysterectomies or simply rendered barren as the result of extensive scarring all because of an abortion to which they never gave informed consent. Abortion is a very dangerous procedure for the mother and child, both physically and psychologically. I, along with many others, became hateful, angry, bitter and resentful after the abortion. I dwelled on the baby's due day for years. My next abortion later was legal and the counseling wasn't any better. I wanted to make up for the first baby I killed. In turn I killed this one also. It saddens me to see how tough and "pro-choice" I became after the second abortion. The first one devastated me and, in order to handle my feelings, I had to become hard and put up a wall. I didn't know how to deal with murder. Can you imagine how it feels years later to see pictures of aborted babies? I thought I would die, seeing what I'd done! I am 100% pro-life now. Women should be told all the facts about abortion. "What you don't know can hurt you!" It kills babies physically and women spiritually. We, the women who have had abortions, know the physical and emotional pain left in abortion's wake. It is and has been our struggle to deal with abortions' aftermath, now it is our struggle to tell others the truth about abortion. Please help us in this effort. Sincerely.</p>
Democrat Result:	<p>Mr. Speaker, 13 years ago today, the right of women to choose abortion in appropriate circumstances was affirmed by the highest court in the land. Abortion is legal, and the overwhelming majority of Americans want to keep it that way. Nevertheless, a vocal minority continues to threaten this fundamental right with a constant barrage of attacks on abortion rights. We must remain ever vigilant against these misguided efforts. As a young prosecutor in New York City charged with investigating and prosecuting illegal abortionists, I witnessed firsthand the tragedy of back alley abortions. No one should want to force anyone's loved ones back to the hands of the kind of butchers who performed abortions in those days before Roe versus Wade. Millions of women risked permanent disability as a consequence of procedures used in the home or in the underground network of back alley abortionists. Hundreds of women died annually as a result of botched abortions. On this day of remembrance, we must recommit ourselves to maintaining the right to choose. We must support the Supreme Court's refusal to weaken Roe versus Wade. And we must restore full abortion rights to those who suffer unfairly under legislated restrictions. Making abortions illegal didn't prevent abortion. It just made them dangerous and often lethal. A woman's right to choose a safe and legal abortion must remain as sacred as any right American citizens enjoy.</p>
Question:	Is climate change real?
Republican Result:	<p>Mr. President, this amendment addresses the language in the preamble which could create the impression that the science is complete and that there is consensus that the Earth is currently threatened by global climate change. This is still an issue in dispute. The science community is engaged in the very active debate on the potential for global climate change and its effects. Next month, the Senate Energy Committee will hold another hearing on global climate change, and this will be an appropriate opportunity to debate the science of it. My amendment reflects my uncertainty, only adding the concept of potential to the threat of climate change.</p>
Democrat Result:	<p>Mr. President, years ago, Will Rogers remarked that everybody complains about the weather but nobody does anything about it. Our task, heading into the 1990s, is to prove Will Rogers wrong. I am here this morning not just to complain about the dangers of global climate change, but also to talk about what we can and must do to reverse mankind's current perilous course. Few dispute the fact that the Earth's climate is changing perceptibly and relatively speaking rapidly. In the past year, the world's consciousness has been jolted by droughts, floods, and heat waves of historic magnitude. We ignore these events and trends at our own peril. On that score, I am personally committed to seeing that this Congress leads the way in improving our understanding of global climate change, and in mobilizing an international response. I know that many of my colleagues share that sense of commitment, and hope that the recently formed Global Warming Task Force will be effective in moving us closer to those goals. Today, we are here to discuss the task force's first initiative: a letter to President Bush urging him to press for action to address global warming at the ongoing economic summit in Paris. The summit provides a critical opportunity to show the American public and the world that the United States will provide responsible leadership in dealing with the threat of global climate change. I join my colleagues in this call to action because the worldwide implications of climate change must transcend traditional political and national rivalries. The bottom line really is that through population growth, industrialization, and energy use, mankind is affecting the way our planet works. In recent years, scientists have observed major changes in the Earth's atmosphere, oceans, and land masses. These changes include rising world temperatures, shifts in climate zones, depletion of the stratospheric ozone layer that shields the planet from harmful radiation, and possible rises in sea levels. If these global changes continue, they will have profound and lasting effects on many aspects of life.</p>

Table IX

EXAMPLES OF QUESTIONS, TOGETHER WITH ONE EXAMPLE OF SEARCH RESULTS FOR BOTH THE REPUBLICANS AND DEMOCRATS.

