A Dataset of General-Purpose Rebuttal
-
IBM Debater(R): Recorded Debating Dataset (V4)
GPR-KB-55 Readme

1 Introduction

This document describes the GPR-KB-55 dataset (Orbach et al., 2019): a collection of 55 general-purpose claim (GP-claim) and rebuttal pairs, and the results of several annotation experiments done using this collection. The annotation experiments were performed on 200 speeches in which an expert human debater is arguing for or against one of 50 controversial topics from iDebate (Mirkin et al., 2018b). Herein we refer to their dataset as iDebate18. The controversial topics in iDebate18 are referred to as *motions*, as in formal parliamentary debates.

The dataset is provided in several `csv` files. This document contains their description, including one example entry from each file.

2 Claims and rebuttals

The file `GPR-KB-55.csv` contains 55 general-purpose claim and rebuttal pairs, and its format is described in Table 1.

The phrasing of the motions in the data of Mirkin et al. (2018b) is often simplified to include an explicit *topic* and *action*. For example, the iDebate motion *This House would introduce goal line technology in football* is simplified to *We should introduce goal line technology*, where the topic is *goal line technology* and the action is *introduce*.

The claim and rebuttal texts were allowed to incorporate the special tokens *<ACTION>* and *<TOPIC>*, which are replaced by the debate's topic and suggested action when applied to a specific motion or speech. For example, in the context of the motion *We should introduce goal line technology*, the claim *<ACTION> <TOPIC> will encourage better choices* is translated to *introducing goal line technology will encourage better choices*.

3 Speech transcripts

GPR-KB-55 contains several annotation layers on top of the recordings collected by Mirkin et al. (2018b). In their data, each recording is accompanied by a manual transcription of the audio, in two formats, one including the time stamps of each word, and another "NLP-friendly" format excluding these time stamps. For convenience, we include the latter for all of their recordings in the file `speech-transcripts.csv`. Its format is described in Table 2. A detailed description of the format of original transcripts is given in Mirkin et al. (2018a).

4 Annotation experiments

Following are four short descriptions of the annotation experiments performed in the paper, along with a description of the files containing the annotation results. Please refer to the paper (Orbach et al., 2019) for a more in-depth discussion of the experiments and their results.

4.1 Cross-topic relevancy

The first annotation experiment aims to establish whether GP-claims indeed attain the desired goal of being applicable to a varied set of topics. For each motion in iDebate18, and for each GP-claim, we asked annotators to decide whether the claim supports the motion, contests it or is not relevant. Annotation was done by 7 experienced annotators, and 5 answers were collected for each question. The results of this annotation experiment are given in the file `GP-claim-to-motion-relevancy.csv`. The format of this file is detailed in Table 3.

A GP-claim is considered *relevant* to a motion when marked as supporting or contesting it by most annotators. The stance of relevant claims towards the motion was determined by majority.

Field	id	claim	rebuttal
Description	A unique id for a GPR-KB unit.	A general-purpose claim (GP-claim).	A rebuttal to the GP-claim.
Example	GPR-KB-1	We need to think about how this affects us right now.	It is necessary to balance short-term and long-term concerns. The long-term effects in this case greatly outweigh the short-term ones.

Table 1: The format of the file `GPR-KB-55.csv`. This file contains claim and rebuttal pairs.

Field	Description	Example
speech-id	A unique id for the speech.	daniel_81_physical-education_con_opening_james
motion-id	A unique id for the motion.	81
speech-to-motion-polarity	The stance of the speech towards the motion. `PRO`: the speech supports the motion. `CON`: the speech contents the motion.	CON
speaker-name	The name of the speaker in the speech.	daniel
responds-to-speaker-name	For speeches contesting the motion (otherwise empty): those speeches recorded in response to a speech in favor of the motion. This field contains the name of the opponent in the supporting speech.	james
speech-text	The text of the manual transcription created for the speech.	Physical education should not be mandatory, and obviously physical education, the reason why ...

Table 2: The format of the file `speech-transcripts.csv`. This file contains several properties for each recorded speech along with the manual transcript of its recorded audio (shortened).

When a relevant claim has an equal number of supporting and contesting answers, its stance is considered undetermined.

4.2 Usage in spoken content

The second annotation experiment aims to determine whether GP-claims are commonly made by people debating the topics to which they are relevant. To assess this, annotators were shown speeches from iDebate18, alongside a matching list of GP-claims determined to be relevant in the previous stage. Specifically, claims annotated as supporting a motion were shown for speeches in which the speaker is arguing in favor of that motion, and vice versa. To allow for a greater number of potential claims, those which at least 2 annotators considered relevant (rather than 3) were included. Claims with an

undetermined stance were excluded. The results of this annotation experiment are given in the file `GP-claims-mentioned-in-speeches.csv`. The format of this file is detailed in Table 4.

A claim is considered *mentioned* in a speech if a majority labeled it as mentioned (i.e. summing up implicit and explicit answer counts). Otherwise it is considered as *not mentioned*. A mentioned claim is *explicit* in the speech if its explicit answers count strictly exceeds its implicit answers count. Otherwise, it is considered *implicit*.

4.3 Where was it said?

The third annotation experiment aims to identify *where* a claim was mentioned in a given speech. The annotation included pairs of a claim and a sentence from one of the speeches, and annotators were asked whether the claim is mentioned in

Field	Description	Example
motion-id	A unique id for the motion.	81
motion-text	The motion under discussion.	Physical education should be mandatory
GPR-KB-id	A unique id for a GPR-KB unit..	GPR-KB-16
GP-claim-text	The general-purpose claim (GP-claim) text.	<ACTION><TOPIC> protects personal choice.
GP-claim-text-for-motion	The GP-claim text after assigning the topic and action of the motion into its <ACTION> and <TOPIC> placeholders.	making physical education mandatory protects personal choice.
supports-count	The number of times the claim was annotated as supporting the motion.	3
contests-count	The number of times the claim was annotated as contesting the motion.	0
irrelevant-count	The number of times the claim was annotated as irrelevant to the motion.	2
is-claim-relevant-to-motion	Is the claim relevant to the motion? (decided by majority). Possible values: `TRUE` or `FALSE`.	TRUE
polarity	The stance of the claim towards the motion (empty for irrelevant claims). Determined by majority. Possible values: `PRO` or `CON`.	PRO

Table 3: The format of the file `GP-claim-to-motion-relevancy.csv`. This file contains the annotation of GP-claim and motion pairs. The GP-claim in the above example was annotated as supports the motion by $3/5$ annotators, and as irrelevant to the motion by $2/5$ annotators. It is therefore considered as relevant to the motion (decided by majority), and its stance towards the motion (polarity) is `PRO`.

the sentence. The results, for GP-claims, are in `GP-claim-to-sentence-matching.csv`. Table 5 contains one entry from this file. Similarly, the results for claims taken from iDebate are present in `idebate-claim-to-sentence-matching.csv`. Table 6 contains one entry from this file.

4.4 Validity of rebuttal arguments

We assessed the effectiveness of the rebuttal arguments using a two-step procedure. First, as in Section 4.2, annotators were shown a speech and a claim, and determined whether the claim is mentioned in the speech. Then, if they marked that claim as mentioned, its pre-written rebuttal was shown, and they were asked whether it is a plausible response to the mentioned claim *in the context of the speech*. We collected 10 answers for each question. The results are given in in `rebuttal-to-speech-matching.csv`. Table 7 contains one entry from this file.

References

Shachar Mirkin, Michal Jacovi, Tamar Lavee, Hong-Kwang Kuo, Samuel Thomas, Leslie Sager, Lili Kotlerman, Elad Venezian, and Noam Slonim. 2018a. A recorded debating dataset. In *Proceedings of LREC*.

Shachar Mirkin, Guy Moshkowich, Matan Orbach, Lili Kotlerman, Yoav Kantor, Tamar Lavee, Michal Jacovi, Yonatan Bilu, Ranit Aharonov, and Noam Slonim. 2018b. Listening comprehension over argumentative content. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 719–724. Association for Computational Linguistics.

Matan Orbach, Yonatan Bilu, Ariel Gera, Yoav Kantor, Lena Dankin, Tamar Lavee, Lili Kotlerman, Shachar Mirkin, Michal Jacovi, Ranit Aharonov, and Noam Slonim. 2019. A dataset of general-purpose rebuttal. *EMNLP*.

Field	Description	Example
motion-id	A unique id for the motion.	81
motion-text	The motion under discussion.	Physical education should be mandatory
speech-id	A unique id for the speech.	daniel_81_physical-education_con_opening_james
GPR-KB-id	A unique id for a GPR-KB unit.	GPR-KB-26
GP-claim-text	The general-purpose claim (GP-claim) text.	<ACTION><TOPIC> will not encourage better choices.
GP-claim-text-for-motion	The GP-claim text after assigning the topic and action of the motion into its <ACTION> and <TOPIC> placeholders.	Making physical education mandatory will not encourage better choices.
explicit-count	The number of times the claim was annotated as mentioned explicitly in the speech.	4
implicit-count	The number of times the claim was annotated as mentioned implicitly in the speech.	3
reject-count	The number of times the claim was annotated as not mentioned in the speech.	0
is-claim-mentioned-in-speech	Is the claim mentioned to the motion? (decided by majority, where both explicit and implicit answers are considered as the claim mentioned in the speech). Possible values: TRUE or FALSE.	TRUE
label	A ternary label specifying whether the claim is mentioned in the speech explicitly, implicitly, or not at all. Possible values: Explicit, Implicit or Negative.	Explicit

Table 4: The format of the file `GP-claims-mentioned-in-speeches.csv`. This file contains the annotation of GP-claim and speech pairs. The GP-claim in the example above was annotated as mentioned explicitly in the speech by $4/7$ annotators, as implicit in the speech by $3/7$, and as irrelevant by $0/7$. It is therefore considered as mentioned in the speech (explicitly).

Field	Description	Example
motion-id	A unique id for the motion.	81
motion-text	The motion under discussion.	Physical education should be mandatory
speech-id	A unique id for the speech.	daniel_81_physical-education_con_opening_james
GPR-KB-id	A unique id for a GPR-KB unit.	GPR-KB-26
GP-claim-text	The general-purpose claim (GP-claim) text.	<ACTION><TOPIC> will not encourage better choices.
GP-claim-text-for-motion	The GP-claim text after assigning the topic and action of the motion into its <ACTION> and <TOPIC> placeholders.	Making physical education mandatory will not encourage better choices.
sentence-index	The zero-based index of the annotated sentence (within the speech).	2
sentence-text	The annotated sentence text.	So, first, why do I think PE class is one of the absolute worst ways in which you absolute eh at which you can get a workout?
positive-count	The number of times the claim was annotated as mentioned in the sentence.	3
negative-count	The number of times the claim was annotated as *not* mentioned in the sentence.	10
is-claim-mentioned-in-sentence	Is the claim mentioned in the sentence? (decided by majority). Possible values: `TRUE` or `FALSE`.	FALSE

Table 5: The format of the file `GP-claim-to-sentence-matching.csv`. This file contains the annotation of GP-claim and sentence pairs. The GP-claim in the above example was annotated as mentioned in the speech by 3/13 annotators and as not mentioned by 10/13. It is therefore considered as not mentioned in the sentence.

Field	Description	Example
motion-id	A unique id for the motion.	81
motion-text	The motion under discussion.	Physical education should be mandatory
speech-id	A unique id for the speech.	james_81_physical-education_con_opening_will
idebate-point-id	A unique id for the idebate point in the data of Mirkin et al. (2018b).	1180
claim-text	The text of the annotated iDebate-claim	Students should have the right to choose which subjects they complete at school
sentence-index	The zero-based index of the annotated sentence (within the speech).	1
sentence-text	The annotated sentence text.	So we're going to explain to you why primary and secondary school students ought not be required to participate in gym or other physical education activities.
positive-count	The number of times the claim was annotated as mentioned in the sentence.	2
negative-count	The number of times the claim was annotated as *not* mentioned in the sentence.	8
is-claim-mentioned-in-sentence	Is the claim mentioned in the sentence? (decided by majority). Possible values: TRUE or FALSE.	FALSE

Table 6: The format of the file `idebate-claim-to-sentence-matching.csv`. This file contains the annotation of iDebate-claim and sentence pairs. The iDebate-claim in the above example was annotated as mentioned in the sentence by $2/10$ annotators and as not mentioned by $8/10$. It is therefore considered as not mentioned in the sentence.

Field	Description	Example
motion-id	A unique id for the motion.	121
motion-text	The motion under discussion.	We should ban boxing
speech-id	A unique id for the speech.	james_121_ban-boxing_con_opening_rachel
GPR-KB-id	A unique id for a GPR-KB unit.	GPR-KB-47
GP-claim-text	The general-purpose claim (GP-claim) text.	These risks exist elsewhere as well.
GP-claim-text-for-motion	The GP-claim text after assigning the topic and action of the motion into its <ACTION> and <TOPIC> placeholders.	These risks exist elsewhere as well.
rebuttal-text	The general-purpose rebuttal text.	Indeed there are some similar cases out there and we should look at those cases as well, and see if they fit the same criteria and should be dealt with accordingly.
rebuttal-text-for-motion	The rebuttal text after assigning the topic and action of the motion into its <ACTION> and <TOPIC> placeholders.	Indeed there are some similar cases out there and we should look at those cases as well, and see if they fit the same criteria and should be dealt with accordingly.
positive-count	The number of times the rebuttal was annotated as being a plausible response to GP-claim the in the context of the speech.	8
negative-count	The number of times the rebuttal was annotated as *not* being a plausible response to GP-claim the in the context of the speech.	2
rebuttal-to-speech-label	Is the rebuttal a plausible response to the GP-claim in the context of the speech? (decided by majority). Possible values: TRUE or FALSE.	TRUE

Table 7: The format of the file `rebuttal-to-speech-matching.csv`. This file contains the annotation of rebuttal and speech pairs. In the above example, the rebuttal was annotated as a plausible response for the speech by $8/10$ annotators, and hence it is considered as a valid response. Here, the GP-claim and its rebuttal do not contain any placeholders (i.e. neither <ACTION> nor <TOPIC>), so the values of the fields `GP-claim-text` and `GP-claim-text-for-motion` are the same (for the same reason, the values of the fields `rebuttal-text` and `rebuttal-text-for-motion` are the same).