

# The SEEMPAD Dataset for Emphatic and Persuasive Argumentation

Elena Cabrio and Serena Villata

Université Côte d'Azur, Inria, CNRS, I3S, France

elena.cabrio@unice.fr ; villata@i3s.unice.fr

## Abstract

**English.** Emotions play an important role in argumentation as humans mix rational and emotional attitudes when they argue with each other to take decisions. The SEEMPAD project aims at investigating the role of emotions in human argumentation. In this paper, we present a resource resulting from two field experiments involving humans in debates, where arguments put forward during such debates are annotated with the emotions felt by the participants. In addition, in the second experiment, one of the debaters plays the role of the *persuader*, to convince the other participants about the goodness of her viewpoint applying different persuasion strategies. To the best of our knowledge, this is the first dataset of arguments annotated with the emotions collected from the participants of a debate, using facial emotion recognition tools.

**Italiano.** *Le emozioni giocano un ruolo importante nell'argomentazione in quanto gli esseri umani uniscono atteggiamenti razionali ad atteggiamenti puramente emotivi quando discutono tra loro per prendere decisioni. Il progetto SEEMPAD si propone di studiare il ruolo delle emozioni nell'argomentazione umana. In questo articolo, presentiamo una risorsa ottenuta tramite due esperimenti empirici che coinvolgono le persone nei dibattiti. Gli argomenti presentati durante tali dibattiti sono annotati con le emozioni provate dai partecipanti nel momento in cui l'argomento viene proposto nella discussione. Inoltre, durante il secondo esperimento, uno dei partecipanti svolge il ruolo di persuasore, al fine di convincere*

*gli altri partecipanti della bontà del suo punto di vista applicando diverse strategie di persuasione. Questa risorsa è peculiare nel suo genere, ed è l'unica a contenere argomenti annotati con le emozioni provate dai partecipanti durante un dibattito (emozioni registrate tramite strumenti di riconoscimento automatico delle emozioni facciali).*

## 1 Introduction

Argumentation in Artificial Intelligence (AI) is defined as a formal framework to support decision making (Rahwan and Simari, 2009; Atkinson et al., 2017). In this context, argumentation is used to achieve the so called *critical thinking*. However, humans are proved to behave differently as they mix rational and *emotional* attitudes.

In order to study the role emotions play in argumentation, we proposed an empirical evaluation of the connection between argumentation and emotions in online debate interactions (Villata et al., 2017; Villata et al., 2018). In particular, in the context of the SEEMPAD project,<sup>1</sup> we designed a field experiment (Villata et al., 2017) with human participants which investigates the correspondences between the arguments and their relations (i.e., support and attack) put forward during a debate, and the emotions detected by facial emotion recognition systems in the debaters. In addition, given the importance of persuasion in argumentation, we also designed a second field experiment (Villata et al., 2018) to study the correlation between the arguments, the relations between them, the emotions detected on the participants, and one of the classical persuasion strategies proposed by Aristotle in rethoric (i.e., *logos*, *ethos*, and *pathos*), played by some participants in the debate to convince the others. In our studies, we selected a behavioral method to extract

<sup>1</sup><https://project.inria.fr/seempad/>

the emotional manifestations. We used a set of webcams (one for each participant in the discussion) whose recordings have been analyzed with the FaceReader software<sup>2</sup> to detect a set of discrete emotions from facial expressions (i.e., happiness, anger, fear, sadness, disgust, and surprise). Participants were placed far from each other, and they were debating through a purely text-based online debate chat (IRC). As a post-processing phase, we aligned the textual arguments the debaters proposed in the chat with the emotions the debaters were feeling while these arguments have been proposed in the debate.

In this paper, we describe the two annotated resources resulting from this post-processing of the data we collected in our two field experiments. Our resource, called the SEEMPAD resource, is composed of two different annotated datasets, one for each of these experiments<sup>3</sup>. The datasets collect all the arguments put forward during the debates. These arguments have been paired by *attack* and *support* relations, as in standard Argument Mining annotations (Cabrio and Villata, 2018; Lippi and Torroni, 2016). Moreover, arguments are annotated with the source of the argument, and the emotional status captured from all the participants, when the arguments are put forward in the debate.

To the best of our knowledge, this is the first argumentation dataset annotated with the emotions captured from the output of facial emotion recognition tools. In addition, this resource can be used both for argument mining tasks (i.e., relation prediction), and for emotion classification in text, where instances of text annotated with the emotions detected on the participants are usually not available. Finally, text-based emotion classification would benefit from the different annotation layers that are present in our dataset.

In the reminder of the paper, Sections 2 and 3 describe the dataset resulting from the two field experiments. Conclusions end the paper.

## 2 Dataset of argument pairs associated with the speaker's emotional status

This section describes the dataset of textual arguments we have created from the debates among the

<sup>2</sup><https://www.noldus.com/human-behavior-research/products/facereader>

<sup>3</sup>Available at <http://project.inria.fr/seempad/datasets/>

participants in Experiment 1 (Villata et al., 2017). The dataset is composed of four main layers: (i) the basic annotation of the arguments<sup>4</sup> proposed in each debate (i.e., the annotation in xml of the debate flow downloaded from the debate platform); (ii) the annotation of the relations of support and attack among the arguments; (iii) starting from the basic annotation of the arguments, the annotation of each argument with the emotions felt by each participant involved in the debate; and (iv) starting from the basic annotation, the opinion of each participant about the debated topic at the beginning, in the middle and at the end of debate is extracted and annotated with its polarity.

The *basic* dataset is composed of 598 different arguments proposed by the participants in 12 different debates. The debated issues and the number of arguments for each debate are reported in Table 1. We selected the topics of the debates among the set of popular discussions addressed in online debate platforms like iDebate<sup>5</sup> and DebateGraph<sup>6</sup>.

In the dataset, each argument proposed in the debate is annotated with an *id*, the participant putting this argument on the table, and the time interval the argument has been proposed.<sup>7</sup> Then, arguments pairs have been annotated with the relation holding between them, i.e., support or attack. For each debate we apply the following procedure, validated in (Cabrio and Villata, 2013):

1. the main issue (i.e., the issue of the debate proposed by the moderator) is considered as the starting argument;
2. each opinion is extracted and considered as an argument;
3. since *attack* and *support* are binary relations, the arguments are coupled with:
  - a the starting argument, or
  - b other arguments in the same discussion to which the most recent argument refers

<sup>4</sup>Note that we annotated as an argument each utterance proposed by the participants in the debate. We did not need then to define guidelines to distinguish arguments or their components in the debate, as it is usually done in the Argument Mining field (Cabrio and Villata, 2018).

<sup>5</sup><http://idebate.org/>

<sup>6</sup>[www.debategraph.org/](http://www.debategraph.org/)

<sup>7</sup>Note that when the argument was put forward by the debater in one single utterance the two time instances (i.e., time-from and time-to) coincide. We used the time interval only when the argument was composed of several separated utterances put forward in the chat across some minutes.

(e.g., when an argument proposed by a certain user supports or attacks an argument previously expressed by another user);

4. the resulting pairs of arguments are then tagged with the appropriate relation, i.e., *attack* or *support*.

To show a step-by-step application of the procedure, let us consider the debated issue *Ban Animal Testing*. At step 1, we consider the issue of the debate proposed by the moderator as the starting argument (a):

**(a)** *The topic of the first debate is that animal testing should be banned.*

Then, at step 2, we extract all the users opinions concerning this issue (both pro and con), e.g., (b), (c) and (d):

**(b)** *I don't think the animal testing should be banned, but researchers should reduce the pain to the animal.*

**(c)** *I totally agree with that.*

**(d)** *I think that using animals for different kind of experience is the only way to test the accuracy of the method or drugs. I cannot see any difference between using animals for this kind of purpose and eating their meat.*

**(e)** *Animals are not able to express the result of the medical treatment but humans can.*

At step 3a we couple the arguments (b) and (d) with the starting issue since they are directly linked with it, and at step 3b we couple argument (c) with argument (b), and argument (e) with argument (d) since they follow one another in the discussion. At step 4, the resulting pairs of arguments are then tagged by one annotator with the appropriate relation, i.e.: **(b) attacks (a)**, **(d) attacks (a)**, **(c) supports (b)** and **(e) attacks (d)**. The reader may argue about the existence of a relation (i.e., a support) between (c) and (d), where (d) supports (c). However, in this case, no relation holds as argument (d) does not really supports argument (c), which basically share the same semantic content of argument (b). Therefore, as no relation holds between (b) and (d), no relation holds either between (c) and (d). We decided to not annotate the

supports/attacks between arguments proposed by the same participant (e.g., situations where participants are contradicting themselves). Note that this does not mean that we assume that such situations do not arise: no restriction was imposed to the participants of the debates, so situations where a participant attacked/supported her own arguments are represented in our dataset. The same annotation task has been independently carried out also by a second annotator on a sample of 100 pairs (randomly extracted), obtaining an IAA of  $\kappa = 0.82$ . The IAA is computed on the assignment of the label “support” or “attack” to the same set of pairs provided to the two annotators.

Topic	#arg	#pair	#att	#sup
BAN ANIMAL TESTING	49	28	18	10
GO NUCLEAR	40	24	15	9
HOUSEWIVES SHOULD BE PAID	42	18	11	7
RELIGION DOES MORE HARM THAN GOOD	46	23	11	12
ADVERTISING IS HARMFUL	71	16	6	10
BULLIES ARE LEGALLY RESPONSIBLE	71	12	3	9
DISTRIBUTE CONDOMS IN SCHOOLS	68	27	11	16
ENCOURAGE FEWER PEOPLE TO GO TO THE UNIVERSITY	55	14	7	7
FEAR GOVERNMENT POWER OVER INTERNET	41	32	18	14
BAN PARTIAL BIRTH ABORTIONS	41	26	15	11
USE RACIAL PROFILING FOR AIRPORT SECURITY	31	10	1	9
CANNABIS SHOULD BE LEGALIZED	43	33	20	13
<b>TOTAL</b>	<b>598</b>	<b>263</b>	<b>136</b>	<b>127</b>

Table 1: Dataset of argument pairs and emotions (#arg: number of arguments, #pairs: number of pairs, #att: number of attacks, #sup: number of supports).

Table 1 reports on the number of arguments and pairs we extracted applying the methodology described before. In total, our dataset contains 598 different arguments and 263 argument pairs (127 expressing the *support* relation and 136 the *attack* relation among the involved arguments).

The dataset resulting from such annotation adds to all previously annotated information (i.e., argument id, the argument’s owner, argument’s relations with the other arguments (attack, support)), the dominant emotion detected using the FaceReader system for each participant in the debate. We investigate the correlation between arguments and emotions in the debates, and a data analysis has been performed to determine the proportions of emotions for all participants. For more details about the correlation between emotions and arguments, we refer the interested reader to (Villata et

al., 2017).

An example, from the debate about the topic “Religion does more harm than good” where arguments are annotated with emotions, is as follows:

```
<argument id="30" debate_id="4" participant="4" time-from="20:43" time-to="20:43" emotion_p1="neutral" emotion_p2="neutral" emotion_p3="neutral" emotion_p4="neutral"> Indeed but there exist some advocates of the devil like Bernard Levi who is decomposing arabic countries. </argument>
```

```
<argument id="31" debate_id="4" participant="1" time-from="20:43" time-to="20:43" emotion_p1="angry" emotion_p2="neutral" emotion_p3="angry" emotion_p4="disgusted">I don't totally agree with you Participant2: science and religion don't explain each other, they tend to explain the world but in two different ways.</argument>
```

In this example, the argument “I don’t totally agree with you Participant2: science and religion don’t explain each other, they tend to explain the world but in two different ways.” is proposed by Participant 4 in the debate, and the emotions resulting from this argument when it has been put forward in the chat are *neutrality* for Participant 2, *anger* for Participant 1 and Participant 3, and *disgust* for Participant 4.

Finally, as an additional annotation layer, for each participant we have selected one argument at the beginning of the debate, one argument in the middle of the discussion, and one argument at the end of the debate. These arguments are then annotated by the annotators with their *sentiment classification* with respect to the issue of the debate: *negative*, *positive*, or *undecided*. The negative sentiment is assigned to an argument when the opinion expressed in such argument is against the debated topic, while the positive sentiment label is assigned when the argument expresses a viewpoint that is in favor of the debated issue. The undecided sentiment is assigned when the argument does not express a precise opinion in favor or against the debated topic. Selected arguments are evaluated as the most representative arguments proposed by each participant to convey her own opinion, in the three distinct moments of the debate. The rationale is that this annotation allows to easily detect when a participant has changed her mind with respect to the debated topic. An example is provided below, where Participant4 starts the debate being undecided and then turns to be positive about ban-

ning partial birth abortions in the middle and at the end of the debate:

```
<arg id="5" participant="4" time-from="20:36" time-to="20:36" polarity="undecided">Description's gruesome but does the fetus fully lives at that point and therefore, conscious of something ? Hard to answer. If yes, I might have an hesitation to accept it. If not, the woman is probably mature enough to judge. </argument>
```

```
<arg id="24" participant="4" time-from="20:46" time-to="20:46" polarity="positive">In the animal world, malformed or sick babies are systematically abandoned. </argument>
```

```
<arg id="38" participant="4" time-from="20:52" time-to="20:52" polarity="positive">Abortion is legal and it doesn't matter much when and how. It's an individual choice for whatever reason it might be. </argument>
```

### 3 Dataset of arguments biased by persuasive strategies

We now describe the corpus of textual arguments, about other discussion topics, collected during Experiment 2 (Villata et al., 2018), in which, together with the participants of the experiment, a *persuader* (PP) was involved to convince the other participants about the goodness of her viewpoint, applying different persuasion strategies. Three kinds of argumentative persuasion exist since Aristotle: *Ethos*, *Logos*, and *Pathos* (Ross and Roberts, 2010; Walton, 2007; Allwood, 2016). *Ethos* deals with the character of the speaker, whose intent is to appear credible. The main influencing factors for *Ethos* encompass elements such as vocabulary, and social aspects like rank or popularity. Additionally, the speaker can use statements to position himself and to reveal social hierarchies. *Logos* is the appeal to logical reason: the speaker wants to present an argument that appears to be sound to the audience. For the argumentation, the focus of interest is on the arguments, the argument schemes, the different forms of proof and the reasoning. *Pathos* encompasses the emotional influence on the audience. If the goal of argumentation is to persuade the audience, then it is necessary to put the audience in the appropriate emotional states. The public speaker has several possibilities to awaken emotions in the audience, like techniques and presentation styles (e.g., storytelling), reducing the ability of the audience to

Dataset						
Topic	Strategy	PP position	#arg	#pair	#att	#sup
SINGLE SEX-SCHOOLS ARE GOOD FOR EDUCATION	Logos	Pro	62	20	12	8
SALE OF HUMAN ORGANS SHOULD BE LEGALIZED	Pathos	Con	37	6	1	5
PARENTS ARE ACCOUNTABLE FOR REFUSING TO VACCINATE THEIR CHILDREN	Logos	Pro	74	17	6	11
THERE SHOULD BE GUN RIGHTS	Ethos	Con	94	24	12	12
GO NUCLEAR	Logos	Pro	87	9	8	1
RELIGION DOES MORE HARM THAN GOOD	Pathos	Con	59	14	6	8
ASSISTED SUICIDE SHOULD BE LEGALIZED	Ethos	Pro	102	29	20	9
USE RACIAL PROFILING - AIRPORT	Logos	Con	34	3	0	3
DEATH PENALTY SHOULD BE SUPPORTED	Pathos	Con	128	27	7	20
TORTURE SHOULD BE USED ON TERRORISTS	Logos	Pro	114	13	2	11
<b>TOTAL</b>			<b>791</b>	<b>162</b>	<b>74</b>	<b>88</b>

Table 2: Dataset of argument pairs and persuasion strategies (PP position: stance of the persuader with respect to the topic of the debate).

be critical or to reason.<sup>8</sup> It is worth noticing that the persuasive strategies are not always mutually exclusive in real world scenario, however, for the sake of simplicity, we consider in this paper that when one of the strategies is applied the other do not hold. In addition to a persuasion strategy, the persuader participated into the debate with a precise stance (pro or con) with respect to the debated issue. Such stance does not change during the debate.

Each argument is annotated with the following elements: debate identifier, argument identifier, participant, and time in which it has been published. For each debate, pairs have been created following the same methodology described in Section 2, and all the relations of attack and support between the arguments proposed by the persuader and those of the other participants are annotated. In this way, we are able to investigate the reactions to PP strategy by tracking the proposed arguments in the debate and the mental engagement index of the other participants. An example of Ethos strategy used against gun rights is the following:

```
<arg id="16" debate_id="8" participant="5"
time="19:46:41"> I've been working in the
educational field in USA, and there no-
thing worse than a kid talking about the
gun of his father. As you cannot say "the
right to carry guns is for people without
a kid only". Then no right at all.
</argument>
```

Table 2 describes this second dataset. Ten topics of debate were selected from highly debated ones in the mentioned online debate platforms, to avoid proposing topics of no interest for the participants. In total, 791 arguments, and 162 arguments pairs (74 linked by an attack relation and 88 by a sup-

port one) were collected and annotated. The number of proposed arguments varies a lot depending on the participants (some were more active, others proposed very few arguments even if solicited), as well as the number of attacks/supports between the arguments. We computed the IAA for the relation annotation task on 1/3 of the pairs of the dataset (54 randomly extracted pairs), obtaining  $\kappa = 0.83$ .

## 4 Conclusions

This paper presented the SEEMPAD resource for empathic and persuasive argumentation. These datasets have been built on the data resulting from two field experiments on humans to assess the impact of emotions during the argumentation in online debates. Several Natural Language Processing tasks can be thought on this dataset. First of all, given that the dataset resulting from the Experiment 1 is a gold standard of arguments annotated with emotions, systems for emotion classification can use it as a benchmark for evaluation. In addition, a comparison of systems' performances on this data compared with the standard dataset for emotion classification would be interesting, given that in SEEMPAD emotions have not been manually annotated but they have been captured from the participants' facial emotion expressions. Second, the dataset from Experiment 2 can be used to address a new task in argument mining, namely persuasive strategy detection, in line with the work of (Duthie and Budzyska, 2018) and (Habernal and Gurevych, 2016).

## References

Jens Allwood. 2016. Argumentation, activity and culture. In *Proceedings of COMMA 2016*, page 3.

<sup>8</sup>For more details, refer to the work of K. Budzyska.

- Katie Atkinson, Pietro Baroni, Massimiliano Giacomin, Anthony Hunter, Henry Prakken, Chris Reed, Guillermo Simari, Matthias Thimm, and Serena Villata. 2017. Towards artificial argumentation. *AI Magazine*, 38(3):25–36.
- Elena Cabrio and Serena Villata. 2013. A natural language bipolar argumentation approach to support users in online debate interactions. *Argument & Computation*, 4(3):209–230.
- Elena Cabrio and Serena Villata. 2018. Five years of argument mining: a data-driven analysis. In *Proc. of IJCAI 2018*, pages 5427–5433.
- Rory Duthie and Katarzyna Budzynska. 2018. A deep modular RNN approach for ethos mining. In *Proc. of IJCAI 2018*, pages 4041–4047.
- Ivan Habernal and Iryna Gurevych. 2016. Which argument is more convincing? analyzing and predicting convincingness of web arguments using bidirectional LSTM. In *Proc. of ACL 2016*.
- Marco Lippi and Paolo Torroni. 2016. Argumentation mining: State of the art and emerging trends. *ACM Trans. Internet Techn.*, 16(2):10:1–10:25.
- Iyad Rahwan and Guillermo R. Simari. 2009. *Argumentation in Artificial Intelligence*. Springer.
- W.D. Ross and W.R. Roberts. 2010. *Rhetoric - Aristotle*. Cosimo Classics Philosophy.
- Serena Villata, Elena Cabrio, Imène Jraïdi, Sahbi Benlamine, Maher Chaouachi, Claude Frasson, and Fabien Gandon. 2017. Emotions and personality traits in argumentation: An empirical evaluation. *Argument & Computation*, 8(1):61–87.
- Serena Villata, Sahbi Benlamine, Elena Cabrio, Claude Frasson, and Fabien Gandon. 2018. Assessing persuasion in argumentation through emotions and mental states. In *Proc. of FLAIRS 2018*, pages 134–139.
- Douglas N. Walton. 2007. *Media argumentation - dialect, persuasion and rhetoric*. Cambridge University Press.