

Towards an Analysis of Argumentation Structure and the Strength of Arguments in News Editorials

Bal Krishna Bal¹ and Patrick Saint-Dizier²

1 INTRODUCTION

Editorials in general fall under persuasive texts. These types of texts intend to persuade the readers over a certain issue or topic. Hence persuasion involves the use of persuasive elements (opinion words and expressions) and facts presented in the form of arguments. A closer look at editorials reveals that they consist of an argumentation structure consisting of an opening statement (also known as the conclusion in argumentation theory [6, 7, 12, 13, 15]), which in turn is supported by other statements (known as the supports in argumentation theory) for or against the conclusion. These supports as well as the conclusion can be either facts or opinions. The underlying supports for or against a conclusion may be further developed, illustrated, justified, elaborated etc. by means of text fragments, also widely known as rhetoric relations [11].

The proposed work is aimed towards analyzing the argumentation structure and the strengths of arguments in news editorials thus determining the persuasiveness inherent in the texts. The result is a discourse analysis of opinions as stated in editorial texts producing a kind of dedicated semantic representation. Ultimately, the analyzed argumentation structure would be used to construct a synthesis of positive and negative arguments on a particular topic from one or several editorials (single or multiple sources) over a common date or a span of time. Such a synthesis can provide a relatively true view of how an event has been perceived by the public in general and is of much interest to journalists, public figures and political analysts. We also will be analyzing the change in opinions taking time as an evaluative factor for change as reported in [8, 9]. In order to automate the analysis as well as the synthesis construction process, we would be developing a computational model that would suggest methods and appropriate techniques. The manually annotated texts and collected editorials would serve as training data and test data respectively for validating the computation model over ideal outputs [1].

Currently, the work is in its preliminary stage, primarily focused towards analyzing the different facets of supports and rhetorical relations required for an adequate semantic-pragmatic analysis of the underlying argumentation structures in editorials. In parallel with the analysis, we are also in the process of specifying tags for annotating editorials in order to establish patterns characteristic to the different facets of support. We noted that the opinions and the argumentation structure in editorials are not so apparent and structured, which makes synthesis construction a challenging task.

¹ Madan Puraskar Pustakalaya, Lalitpur, Patan Dhoka, Nepal, email: bal@mpp.org.np

² IRIT, 118 Narbonne 31062 Toulouse, France, email: stdizier@irit.fr

2 PERSUASIVE TEXTS AND THE UNDERLYING ARGUMENTATION STRUCTURE

In order to make the text convincing or persuasive, the common practice is to follow one or more of the following strategies [14]:

1. **Use of logical and cause effective connectives like *however, so, and, although* etc.** Example - However, if we compare the present situation with the period before 2006, we have definitely come a long way.
2. **Trying to make opinions sound like facts.** Example - The year also saw the height of anarchy, impunity and lawlessness.
3. **Use of powerful adverbs and adjectives.** Example - The Post believes that the long awaited and ever elusive CA elections will take place this year, and that the country will take the course to sustainable peace and development.
4. **Use of words like *surely, obviously, of course, definitely* etc.** Example - If the Maoists do not run away from the elections, if the recently formed and old parties in the terai live up to the promises to allow the CA elections to take place, and if the government deals with other problems appropriately, the year 2008 will definitely herald the beginning of a new Nepal.

Source: Editorial - "Year of hope", The Kathmandu Post, December 28, 2007 (<http://ekantipur.com/ktmpost.php>)

3 SUPPORT AND RHETORICAL RELATIONS IN THE ARGUMENTATION STRUCTURE OF EDITORIALS

In our semantic and pragmatic representation of editorials, the root node is a conclusion. This conclusion bears a polarity: positive, negative or neutral. It has also a date and a source. Obviously, this polarity is either inherent through explicit linguistic marks, or needs to be deduced. In the latter case, depending on the view point of the reader, the polarity may vary.

Next, the root node is associated with one or more supports. The support relation, we define, consists of six fields that represent the facets of the support. These are represented by means of attribute-value pairs. These facets are:

- The date the support has been uttered.
- The source (name of newspaper, name of utterer if it is a reported opinion).
- The orientation of the support, namely for or against the conclusion. At the moment, we are not completely satisfied with the plain

for-against dichotomy, and hence are exploring deeper modes of classification, to better capture the notion of orientation.

- The reporting level included in the support, if any. The higher, and by-default level, is the say level, characterized by verbs like: note, notice, remark, etc. with little commitment on the part of the utterer (with attribute value commitment: low), the next level, with a stronger commitment is characterized by report verbs (inform, support, reveal, pretend, etc., commitment: high). Finally, a number of supports showing weaker commitments, characterized by the epistemic or modal verbs such as: think, believe, assume, etc. (commitment: modal (Verb)). At this level, we do not go into further classification details since these verbs entail quite debatable modal scales.
- The conditional level, when a condition introduces the support, making it relative to another statement in the conditional part. Values for that feature are conditional: yes or no.
- The strength of the argument. This value is further divided into several fields: (1) the visible and direct strength as characterized by the polarity and the inherent strength of the terms used (direct-strength: low, average, high), (2) the relative strength, elaborated from the strength of the other supports; this allows us to take into account the personal style of the writer (relative-strength: low, average, high), and, (3) the persuasion effect that takes into account social, stylistic, contextual as well as typographical aspects, in addition to level 1.

Next talking about rhetorical relations, editorials subsume a large variety of rhetorical relations. Some are more central than others. For our purpose, we have identified the following subset of relations which really do play a crucial role:

- Exemplification: illustrates a support, while giving it a higher strength and persuasion effect.
- Contrast: relates two supports A and B, where A and B are both true while partly contradicting each other. They are in general linked by connectors such as nevertheless, although, but, even if, etc.
- Discourse Frame: introduces a factual statement which indicates the environment and the scope of the conclusion (time, facts, etc.), without being a support (Last year, Nepal was declared a Federal Democratic Republic).
- Justification: where B gives reasons and explains A, this relation is stronger than the explanation relation.
- Elaboration: where B is an elaboration of A if it develops or describes a part of A.
- Paraphrase: which is just another way of saying a support or a conclusion, adding strength to it.
- Cause-effect: established a causal relation between supports.
- Result: where B results at least partly, or indirectly, from A.
- Explanation: where B is an explanation for A if it indicates the reasons for A, in a quite neutral way.
- Reinforcement: where B gives a stronger weight to A by its contents. It is stronger than an elaboration, an exemplification or an explanation. In general it contains specific marks, related e.g. to confirmation, enforcement, etc.

Below, we present an example of the argumentation structure from our corpus of editorials.

Conclusion: (<Date:2007-12-28>, <Source:KTMPOST>, <Orientation:Positive>, <Strength:High>)
[CA elections] will take place in 2008.

Support: (<ID:1>, <Date:2007-12-28>, <Source:KTMPOST>,

<Orientation:Positive,Support Type:For>, <Strength:Low>)

The Post believes that the long awaited and ever elusive [CA elections] will take place this year.

Rhetorical_relation:Justification(1,2)

Support: (<ID:2>, <Date:2007-12-28>, <Source:KTMPOST>, <Orientation:Positive,Support Type:Conditional,For>, <Strength:Low>)

If we behave responsibly, we will be able to hold the [CA elections].

Rhetorical_relation:Justification(1,3)

Support: (<ID:3>, <Date:2007-12-28>, <Source:KTMPOST>, <Orientation:Positive,Support Type:Conditional,For>, <Strength:Low>)

If the Maoists do not run away from elections, if the recently formed and old parties of the terai live up to the promises to allow [elections] happen...

In the example above, the conclusion is characterized by a vector that contains id, date, source, orientation and strength. The conclusion is followed by supports and rhetorical relations. The latter establishes additional information on supports. This follows that unlike a linear model of argumentation as discussed in [10], the argumentation structure in editorials are more of a connected graph model. The description of supports is also done in the same way as in the case of the conclusion. We put the referential expression inside square brackets, which binds the supports to the event reported in the conclusion. It should be noted that **[CA elections]** and **[elections]** are the referential expressions in the example above. Similarly, the underlined text portions above are the opinion anchors, i.e., those terms that a priori mark the statement as an opinion. For the strength, we are currently only considering the attribute, direct-strength. Other two attributes would be gradually incorporated.

As noted in the introduction section, the argumentation structure in editorials generally consists of two parts, viz., conclusion and the supportive arguments. However, we have observed that in some of the editorials, the conclusion and the supportive arguments interchange positions. Similarly, supportive arguments are found to be linked or elaborated by rhetorical relations. For instance, in the example above, it is to be noted that there is a rhetorical relation of type - justification in between supports with ids 1 and 2 as well as supports with ids 1 and 3.

4 LINGUISTIC FRAMEWORK AND THE DISTINCTION BETWEEN FACTS AND OPINIONS

Since editorials are usually a mix of facts and opinions, there is a clear need to make a distinction between them. Opinions often express an attitude towards something. This can be a judgment, a view or a conclusion or even an opinion about opinion(s). Different approaches have been suggested to distinguish facts from opinions [3, 4, 5, 16]. Generally, facts are characteristic for the presence of certain verbs like *declare* and different tense and number forms of the verb *be* etc. Moreover, statements interpreted as facts are generally accompanied by some reliable authority providing the evidence of the claim, e.g.:

Fact: Both the two dates announced for the constituent assembly (CA) elections came and went without the vote taking place.

Reliable authority: *Election Commission for CA elections 2007.*

Fact: We have fewer people getting killed every day.

Reliable authority: *Nepal Police Department of Crime and Investigation. (December 2007)*

Opinions, on the other hand, are characterized by the evaluative expressions of various sorts such as the following [3]):

- Presence of evaluative adverbs and adjectives in sentences - *ugly* and *disgusting*.
- Expressions denoting doubt and probability - *may be, possibly, probably, perhaps, may, could* etc.
- Presence of epistemic expressions - *I think, I believe, I feel, In my opinion* etc.

It is obvious that the distinction between the two is not always straightforward. Facts could well be opinions in disguise and, in such cases, the intention of the author as well as the reliability of information needs to be verified. In order to make a finer distinction between facts and opinions and within opinions themselves, opinions are proposed for gradation as shown below:

Opinion type	Global definition
Hypothesis statements	Explains an observation.
Theory statements	Widely believed explanation
Assumptive statements	Improvable predictions.
Value statements	Claims based on personal beliefs.
Exaggerated statements	Intended to sway readers.
Attitude statements	Based on implied belief system.

5 STRENGTH OF ARGUMENTS

The work reported in [5] employ gradability of adjectives as a major factor for determining the strength of opinions or grades of subjective expressions in sentences. We extend this idea to arguments for which we formulate attribute-value pairs depending upon whether the argument falls under facts or opinions. Opinion arguments are further subcategorized into three groups, *High, Medium* and *Low* depending upon the different forms of opinion adjectives they contain. For instance, if the *Argument type* is a *Fact*, then its strength would be *High*. Similarly, any opinion expression like *large audience* would receive the value of strength equal to *Low*, whereas (*larger audience* and *the largest audience* would receive *Medium* and *High* respectively. However, this is just a general purpose scheme and more precise specifications are necessary for dealing with more complex expressions, whose strength cannot alone be determined by the proposed technique.

To make the determination of the strength of opinions more precise and accurate, we categorize opinion words and expressions collected from our corpus into prototypically positive and negative sets as discussed in [4,16]. Moreover, as part of the categorization, we propose to group semantically similar members under different subsets within the bigger sets. The smaller subsets would bear the name from one of the members belonging to that particular subset. For instance, the negative set {*poor*(adj), *miserable*(adj), *trouble*(n), *troublesome*(n), *troubling*(adj)} could be named as the *miserable* set. Such a subcategorization would ease in the clear cut determination of the polarity of opinion words and expressions.

Next the categorized sets would be further subject to subcategorization on the basis of the strengths as exhibited by the members of the sets. For this purpose, we further split the strength attribute into three sub-attributes, namely *direct-strength, relative-strength* and *persuasion strength*, which respectively take one of the values low,

medium or high. The entries of the subsets and consequently the sets are then arranged in the same manner as in the case of operators separated by commas in a precedence table. In (Bal & Saint-Dizier, Forthcoming), we provide a detailed specification of this process. Our approach discussed above is somewhat different from [2], which takes into account the predefined values and beliefs of the readers as a crucial factor in persuasion.

6 TEXT COLLECTION AND ANNOTATION

Editorials have been collected from at least three different sources. The collected texts serve as a corpus for our research work. The editorials represent a common theme - *Soci-political* and subtheme *Peace and stability* and are taken from different dates towards the end of the year 2007 and the beginning of 2008 amounting a total of 300 plus text files, with a total of approximately 6000 sentences and an average of 20 sentences per editorial. The texts are taken respectively from *The Kathmandu Post Daily*, <http://ekantipur.com/ktmpost.php>, *The Nepali Times Weekly*, <http://nepalitimes.com.np> and *The Spotlight Weekly*, <http://nepalnews.com/spotlight.php>. Two annotators having a fairly good understanding of the English language have been involved in the annotation work. The annotators have been assigned the same texts to see how semantic annotations can differ among annotators. Results have shown that the difficulties in the manual annotation exist at two levels, the first one in determining the orientation of polarity of words or expressions and the second one in evaluating their strengths for the three different strength attributes - *direct-strength, relative-strength* and *persuasion-strength*. Wherever the annotators have confusions about providing one particular value, they have been advised to provide multiple values separated by commas.

For the annotation purpose, we have developed a semantic tagset, subject to further extension or modification in future. The current tagset can be represented as a list of parameters and their possible values as shown below:

Parameters	Possible values
argument_type	Support, Conclusion, Rhetorical_relation
expression_type	Fact, Opinion, Undefined
fact_authority	Yes, No
opinion_orientation	Positive, Negative, Neutral
orientation_support	For, Against
id	Id number of the support
date	Date of publication of the editorial
source	Source or name of the newspaper
commitment	Modal, Low, High
conditional	Yes, No
direct-strength	Low, Average, High
relative-strength	Low, Average, High
persuasion-effect	Low, Average, High
rhetoric_relation_type	Exemplification, Contrast
	Discourse Frame, Justification
	Elaboration, Paraphrase
	Cause-effect, Result
	Explanation, Reinforcement

The tagset has been used to annotate the texts in XML format for outlining the argumentation structure and strength of the argument. Below, we provide a sample of the annotated text.

<Arguments>

<Conclusion date="Jan 05,2006" source="The Kathmandu Post" orientation="Negative">
The actions of the power centers seems to be heading to confrontational poll and politics.
</Conclusion>

<Support id="1" date="Jan 05,2006" source="The Kathmandu Post" orientation_support="For" commitment="Modal" conditional="No" direct-strength="High" relative-strength="High" persuasion-effect="High">
It is utter naivety on the part of the royal government to believe that the municipal polls would fix all the problems of the country.
</Support>

<Rhetoric_relation type="Elaboration" has_relation_to_support="1">
In fact,it will neither lessen the woes of the country, nor will it give any legitimacy to the autocratic monarchy.
</Rhetoric_relation>

<Support id="2" date="Jan 05,2006" source="The Kathmandu Post" orientation_support="For" commitment="High" conditional="No" direct-strength="High" relative-strength="High" persuasion-effect="High">
The polls will only widen the rift between the political parties and the royalists, and it is certain to invite more bloodshed.
</Support>

<Support id="3" date="Jan 05, 2006" source="The Kathmandu Post" orientation_support="For" commitment="High" conditional="No" direct-strength="High" relative-strength="High" persuasion-effect="High">
The government has remained obdurate,showing no intention to reciprocate the peace initiative taken by the Maoist rebels.
</Support>

<Rhetoric_relation type="Elaboration" has_relation_to_support="3">
The unilateral cease fire announced by the Maoists expires today.
</Rhetoric_relation>

<Rhetoric_relation type="Elaboration" has_relation_to_support="3">
It is unfortunate that the royal government is not persuaded by any level of persuasion to work for peace.
</Rhetoric_relation>

<Support id="4" date="Jan 06,2006" source="The Kathmandu Post" orientation_support="Against" commitment="High" conditional="No" direct-strength="High" relative-strength="High" persuasion-effect="High">
Currently the seven-party alliance is dead against the holding of the municipal polls scheduled for February 8.
</Support>

</Arguments>

ACKNOWLEDGEMENTS

I am very grateful to Professor Patrick Hall,my PhD co-supervisor,for his continuous support and guidance.I am also grateful to my institution,Madan Puraskar Pustakalaya for all the support and platform provided to me.This work was partly supported by the French Stic-Asia programme.

REFERENCES

- [1] Bal B.K.,”Towards an Analysis of Opinions in News Editorials: How Positive was the year?,”in Proceedings of the 8th International Conference in Computational Semantics,Tilburg,Netherlands,2009, pp. 260-263.
- [2] Bench-Capon T.J.M.,”Agreeing to Differ:Modelling Persuasive Dialogue Between Parties Without a Consensus About Values,”*Informal Logic*,vol.22,no.3,pp.231-45, 2002.
- [3] Dunworth K.(2008)UniEnglish reading:distinguishing facts from opinions. [Online].HYPERLINK
”http://unienglish.curtin.edu.au/local/docs/RW_facts_opinions.pdf”
- [4] Esuli A.and Sebastiani F.,”Determining term subjectivity and term orientation for opinion mining,” in Proceedings of EAACL-06, 11th Conference of the European Chapter of the Association for Computational Linguistics,Trento,IT.,2006.
- [5] Hatzivassiloglou V. and Wiebe J. M.,”Effects of adjective orientation and gradability on sentence subjectivity.,” in Proceedings of the 18th Conference on Computational Linguistics - Volume 1,Saarbrcken,Germany,2000,pp.299-305.
- [6] Hoffmann M.H.G.,”Logical argument mapping: a cognitive-change-based method for building common ground,”in Proceedings of the 2nd international conference on Pragmatic web,Tilburg,The Netherlands,2007,pp. 41-47.
- [7] Kirschner P.A.,Shum S.J.B.,and Carr C.S.,Visualizing argumentation:software tools for collaborative and educational sense-making.:Springer-Verlag,2003.
- [8] Ku L.-W.,Lee L.-Y.,and Wu T.-H.,”Major topic detection and its application to opinion summarization.,” in SIGIR 2005,pp.627-628.
- [9] Ku L.-W.,Liang Y.-T.,and Chen H.-H.,”Opinion extraction,summarization and tracking in news and blog Corpora.,”in AAAI-2006 Spring Symposium on Computational Approaches to Analyzing Weblogs, AAAI Technical Report.,2006.
- [10] Manfred S.and Sauer mann A.,”Linearization of arguments in commentary text,” in Multidisciplinary Approaches to Discourse 2008 - MAD 08, 20-23 February 2008, Lysebu, Oslo, 2008.
- [11] Marcu D.,”The rhetorical parsing of natural language texts,”in Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics,Madrid,Spain,1997,pp.96-103.
- [12] Moens M.-F.,Boiy E.,Palau R. M.,and Reed C.,”Automatic detection of arguments in legal texts,”in Proceedings of the 11th international conference on Artificial intelligence and law,Stanford,California,2007,pp. 225-230.
- [13] Prakken H.,Reed C.,and Walton D.,”Argumentation schemes and generalisations in reasoning about evidence,”in Proceedings of the 9th international conference on Artificial intelligence and law,Scotland,United Kingdom,2003,pp.32-41.
- [14] Primary Resources:English:Text Level:Persuasive Writing. [Online].HYPERLINK <http://www.primaryresources.co.uk/english/englishD10.htm#persuasive>
- [15] Walton D.N.,Argumentation Schemes for Presumptive Reasoning.Mahwah,NJ:Lawrence Erlbaum Associates, 1996.
- [16] Wilson T.,Wiebe J.,and Hoffman P.,”Recognizing contextual polarity in phrase-level sentiment analysis.,”in Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, Vancouver,British Columbia,Canada,2005.