

KATHMANDU UNIVERSITY
SCHOOL OF ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

DISSERTATION



COMPUTATIONAL LINGUISTIC MODEL FOR ANALYZING
OPINIONATED TEXTS

In Partial Fulfillment of the Requirements for the Doctor of Philosophy Degree in
Computer Science & Engineering

by

Bal Krishna Bal

January, 2015

© 2015, Bal Krishna Bal

ABSTRACT

Editorials and opinion articles in general express their views or opinions about some particular event or happening. These views or opinions are meant to influence or convince the readers and hence may resort to any means of expression in the text in order to make the text sound convincing. In order to support one's thesis or statement, the author may produce facts or mere opinions in disguise of facts. It can also involve exaggerations, sarcasm and irony, biases etc. This makes analyzing the opinions and the argumentation structure in editorials quite a challenging task. Fully understanding the intended meanings conveyed by editorials is also an equally challenging task as knowledge of previous discourse and contexts is required. A synthesis of consolidated information from multiple sources could be a step towards getting a balanced account of events. This research aims to build the technology required for such a synthesis.

Preliminary knowledge about the research problem is gathered via an extensive literature review from a number of interrelated disciplines and sub-disciplines of Artificial Intelligence and Computational Linguistics thus building on partial solutions available. Based on the knowledge gathered from the literature review, specifications of the linguistic features required for analyzing opinions and arguments in opinionated texts are developed. The specifications developed are then validated via a manual analysis of real opinionated texts, as a result of which a semantic tagset is developed for annotating the opinions and argumentation elements in opinionated texts. The viability and usefulness of the semantic tagset is then studied through an inter-annotator agreement study for annotating real opinionated texts.

Existing research problems in Opinion Mining are still too focused in the the lower levels, viz., lexical, phrase and sentence levels. Even in these lower levels, issues regarding inadequate lexicon coverage and poor adaptability of the lexicons in cross-domain situations prevail. There are increasing trends of the use of machine learning approaches for solving one or more of these tasks but as the scope of the analysis and synthesis gets wider together with the involvement of

multiple attributes for the task, the role of lexicon and dictionary based approaches along with the data and patterns from the corpus becomes quite prominent. The given work extends the analysis of opinions in the paragraph level by borrowing the concepts like "opening-statement-thesis" and "supports" from argumentation theory and "rhetorical relations" from rhetorical structure theory and discourse analysis. A synthesis of opinions and arguments from multiple sources is conducted based on the analyses of opinions in the lexical, phrase, sentence and paragraph levels in the form of annotations.

This research work conducts a feasibility study of a complete system for analysis and synthesis of opinions and arguments via a Computational Linguistic Model that suggests appropriate techniques and methods for consolidating material from a number of editorials on a particular topic, taking into account the opinion orientations in these editorials. The resources required for the model are developed, which involves downloading and compiling editorials from a number of online news portals, constructing linguistic resources (lexicons, opinion patterns and contextual rules, semantic tagset), manually and automatically annotating the editorials with some crucial information like opinion orientation, opinion strength, persuasion effect, commitment level etc. Results have shown that the task can be accomplished with reasonable accuracies. As the resources get more fine-tuned and attain wider coverage, the accuracies are expected to increase further. The development of a complete system and its evaluation has been left as future extensions to this work.

LIST OF ABBREVIATIONS AND ACRONYMS

DCU - Discourse Constituent Units

DM - Discourse Marker

EM – Expectation-Maximization

ID – Identity

ILPRL – Information and Language Processing Research Lab

IR - Information Retrieval

IRIT – Institut de Recherche en Informatique de Toulouse

NATO – North Atlantic Treaty Organization

NLP - Natural Language Processing

NLTK - Natural Language Toolkit

ORS - Overall Relatedness Score

PMI - Pointwise Mutual Information

POS - Part-of-Speech

RS - Relatedness Score

RST - Rhetorical Structure Theory

SO - Semantic Orientation

SVM - Support Vector Machines

URL - Uniform Resource Locator

XML - Extended Markup Language

YCL - Young Communist League

UGC – University Grants Commission

ACKNOWLEDGEMENT

This work merits the acknowledgement to many individuals and organizations. First of all, I would like to express my deepest gratitude to my supervisor Professor Patrick Saint-Dizier, who has been supporting me immensely right from the beginning. He was so kind to facilitate my enrollment in the PhD program at the University de Toulouse, University – Toulouse III, Paul Sabatier. Besides, he also helped to arrange support for attending a few summer schools in the region and periodic research visits to IRIT labs, Toulouse in the early years of the PhD. Furthermore, he has been a pro-active supervisor and guide, always prompt with advice and suggestions over emails despite his busy schedule. My other supervisor, Professor Patrick A.V. Hall has provided extraordinary help and guidance to give this thesis the current shape. He has been a very attentive reader of my drafts and has provided feedbacks to even minute details in order to improve the overall thesis. He also has been always eager to help me in every possible way, sometimes pointing me to useful literature and at other times, sending me books all the way from the U.K. I am very thankful to him. He has been more of a guardian and a well-wisher besides the supervisor to this work.

I am equally thankful to Professor Bhola Thapa, the current Registrar of Kathmandu University who has been very supportive to me in all respects, be it for my PhD studies and the transfer to Kathmandu University or in the professional engagements. I should also not miss to thank Mr. Amar Gurung, my friend and a senior colleague back at Madan Puraskar Pustakalya, Nepal – my former work place prior to joining Kathmandu University. He has been a constant source of inspiration to keep motivated with my PhD studies both during hard and easy times.

I would like to express my sincere thanks to MS by Research students, Mr. Rohit Man Amatya and Mr. Chandan Prasad Gupta and the interns – Mr. Shishir Bhattarai, Mr. Shalil Awaley,

Mr. Bishal Kumar Shrestha, Mr. Ashim Neupane and Mr. Anmol Gupta for working in the Opinion Mining Project at the Information and Language Processing Research Lab (ILPRL) and helping me with the technical aspects of the thesis. I would also like to thank the University Grants Commission (UGC), Nepal for the support to the Opinion Mining Project, which has helped to establish the Opinion Mining Research Group at ILPRL.

My family members, especially, my father and mum, my dear wife Radhika and my two children, Rajab and Rajani all deserve special mention. Without their understanding, help and tremendous support, it would not have been simply possible to complete this work. I am also thankful to my relatives, friends and well-wishers who have helped me work with the thesis directly or indirectly.

Last but not the least, I would like to thank my colleagues at the Department of Computer Science and Engineering, Kathmandu University for every support provided to me during the study.

Contents

DECLARATION	iv
ABSTRACT.....	v
LIST OF ABBREVIATIONS AND ACRONYMS	vii
ACKNOWLEDGEMENT	viii
Chapter 1 INTRODUCTION.....	18
1.1. Motivations for this study.....	18
1.2. Objectives and goals.....	21
1.3. Research method.....	22
1.4. Outline of the dissertation.....	24
Chapter 2 LITERATURE REVIEW	26
2.1. Opinion Mining and Related Problems.....	26
2.2. Editorials and Persuasive Writing.....	35
2.3. Argumentation Theory and Structure	36
2.4. Discourse Analysis.....	41
2.5. Summary	44
Chapter 3 CONSTRUCTING THE CORPUS.....	46

3.1. Motivations behind building a Corpus.....	46
3.2. Argumentation Structure in Texts from the Corpus	51
3.3. Devising a semantic tagset.....	53
3.4. Manual annotation of the texts.....	60
3.5. Conclusion	68
Chapter 4 LINGUISTIC RESOURCES FOR AUTOMATIC ANALYSIS.....	69
4.1. Sentiment/Polarity Lexicon	69
4.2. Opinion Bearing Terms and Patterns from the Corpus.....	75
4.3. Intensifier/Pre-modifier Lexicon	77
4.4. Report and Modal Verbs Lexicon.....	81
4.5. Discourse Marker & Rhetorical Relations Lexicon.....	84
4.6. Automatic Annotation Tool	86
4.7. Conclusion	89
Chapter 5 SYNTHESIS OF OPINIONS AND ARGUMENTS	91
5.1. Retrieving Related Texts from Opinionated Texts	92
5.2. Deriving Related and Domain-Specific Terms from the Corpus.....	96
5.3. Experimental Results on Retrieval of texts.....	100
5.4. Evaluation of the Retrieved Results.....	108

5.5. Extracting the Information of the Related Texts from the Annotated Texts	113
5.6. Constructing a Synthesis of Extracted Information from Annotated Texts.....	116
5.7. Conclusion	118
Chapter 6 COMPUTATIONAL LINGUISTIC MODEL.....	119
6.1. The Components of the Model	119
6.2. Fetch opinionated texts	120
6.3. Pre-process texts	121
6.4. Annotate texts	121
6.5. Retrieve the related texts.....	121
6.6. Extract annotations for related texts	122
6.7. Produce a synthesis of opinions and arguments	122
6.8. Conclusion	122
Chapter 7 Conclusions and Recommendations for Future works.....	123
7.1. Contributions to Knowledge.....	123
7.2. Extensions for Future works	125
References.....	127
Bibliography	137
Appendix A Publications	147

Appendix B	Sample of editorials and blogs from the corpus	148
Appendix C	Sample of manually annotated editorials and blogs from the corpus	166
Appendix D	Top 200 most frequent discourse markers from the corpus	218
Appendix E	Sentiment/Polarity Lexicon.....	220
Appendix F	Top 200 frequent domain terms (Nepali Times & The Kathmandu Post, 2007-2008)	233
Appendix G	Extracted bi-grams from the corpus	235
Appendix H	List of English stop words	287

List of Figures

Figure 2-1: An Example Summary	32
Figure 2-2: Abstract Syntax Representation of Toulmin’s Model.....	38
Figure 2-3: Structure of editorials in Abstract Syntax Representation	40
Figure 3-1: Sample of the manual annotation.....	61
Figure 4-1: Excerpt of the analyzed text from the corpus	73
Figure 4-2 : Excerpt of the analyzed text for the strength of opinions and commitment level	84
Figure 4-3: Excerpt of the analyzed text for Discourse Marker and Rhetorical Relations	86
Figure 4-4: High-level system architecture of the annotation tool	87
Figure 5-1: Thesis statement – A year of hope, KTMPOST, Jan 02, 2008	93
Figure 5-2: Text/Paragraph Sample 1 – A year of hope, KTMPOST, Jan 02, 2008	94
Figure 5-3: Text/Paragraph Sample 2 – Peace train, Nepali Times, Dec 28, 2007 – Jan 03, 2008.....	94
Figure 5-4: Relative frequencies of co-occurrences of terms in the corpus	98
Figure 5-5: Bar chart diagram of test cases versus mean accuracies	112
Figure 5-6: Sample of the related texts with respect to the thesis statement	114
Figure 5-7: Sample annotated text for the related text – Year of hope, KTMPOST	115
Figure 5-8: Sample annotated text for the related text – Peace train, Nepali Times	115

Figure 5-9: Sample of synthesis of extracted texts with annotated attributes.....	116
Figure 5-10 : Diagrammatic representation of the synthesis of extracted texts.....	117
Figure 6-1: High level architecture of the computational model.....	120

List of Tables

Table 2-1: Patterns of POS tags for extracting two-word phrases	29
Table 3-1: Information of the corpus	47
Table 3-2: Top 50 frequent Discourse Markers in the corpus	49
Table 3-3: Semantic tagset.....	55
Table 3-4: Illustration of the annotation process	61
Table 3-5: Inter-annotator agreement for the attribute “opinion_orientation” on a text sample	64
Table 3-6: Pair-wise inter-annotator agreements in the form of a contingency table.....	65
Table 3-7: Observed agreements and agreements by chance (contingency table with proportions).....	66
Table 3-8: Inter-annotator agreement scores for some major tagging attributes	66
Table 4-1: A sample of the Sentiment/Polarity Lexicon	71
Table 4-2: Sample of the extracted bigrams from the corpus.....	75
Table 4-3: Sample of the Positive and Negative Opinion Patterns from the Corpus	76
Table 4-4: Intensifier lexicon.....	79
Table 4-5: Pre-modifier lexicon.....	80
Table 4-6: Modal verbs lexicon	82
Table 4-7: Reporting verbs lexicon.....	82

Table 4-8: Discourse markers and rhetorical relations lexicon	84
Table 4-9: Accuracy scores of the automatic annotation tool.....	88
Table 4-10: Accuracy scores of the tool with the baseline and extended linguistic resources	89
Table 5-1: Relative co-occurrence frequencies of the terms in the corpus.....	97
Table 5-2: Top 20 Domain-Specific Terms from the Corpus for the year 2007-2008	99
Table 5-3: Domain-Specific Terms (Phrases) from the Corpus for the year 2007-2008..	99
Table 5-4: Example queries	100
Table 5-5: Computing the relatedness scores of query terms and overall relatedness scores for text/paragraph.....	101
Table 5-6: Texts/Paragraphs with their overall relatedness scores in decreasing order .	104
Table 5-7: Test cases and query parameters	109
Table 5-8: Test cases and their mean accuracies	111

CHAPTER 1 INTRODUCTION

1.1. Motivations for this study

With the increasing interest of the general public towards socio-political happenings, there is a growing need these days for an automatic process to read and analyze the different opinions and arguments on a particular event published by the media in the form of editorials and opinionated texts (from now on opinionated texts, for short, wherever the term ‘editorials’ and ‘opinionated texts’ come together). Such an analysis would help to understand how a particular event has been perceived by different media sources. Hence, it is of primary interest to journalists, public figures and political analysts. A very close example of this form of analysis is available in (Nepal Monitor: The National Online Journal), where the top 10 news stories of Nepal for the year 2009 are briefly presented¹. With regards to the analysis, a more meaningful and useful form of the presentation would have been a synthesis of opinions and arguments on a given topic, taking into account the orientation of opinions (Positive, Negative and Neutral), strength of opinions, persuasion effect and commitment levels etc. together with some useful information like source, date, Uniform Resource Locator (URL) clearly indicated. Such a provision would help the readers get a vivid picture of the happenings or events, yet not necessarily forcing them to go through all the news sources, thereby also potentially allowing to judge the different degrees of biases and prejudices evident in them. Even better would have been the case, if there were a mechanism to track changes in opinions across opinionated texts over a common topic with time.

¹ http://www.nepalmonitor.com/2009/12/nepal_2009_the_top_10_news_stories-print.html

However, it is easier said than done. These problems are quite difficult for the humans, let alone the machine. First and foremost, the discourse of arguments in such texts are not necessarily confined just in the sentence level but span to much higher levels, i.e., one to several paragraphs. Since the views or opinions in opinionated texts are meant to influence or convince the readers, they may resort to any means of expressions in the text in order to make the text sound convincing. In order to support one's viewpoint, the author may produce facts or mere opinions in disguise of facts. The views also can involve exaggerations, sarcasm and irony, biases etc. This makes analyzing the opinions and arguments in opinionated texts a challenging task. Fully understanding the intended meanings conveyed by opinionated texts is also an equally challenging task as knowledge of previous discourse and contexts is required. In his *Sentiment Analysis and Opinion Mining* book, Liu (2012, p.88) notes the following on the analysis of opinions in political texts:

“So far, the research community has mainly focused on opinions about electronic products, hotels and restaurants. These domains are easier (although not easy) and reasonably good accuracies can be achieved if one can focus on each domain and take care of its special cases. When one moves to other domains, e.g., mattress and paint, the situations get considerably harder because in these domains many factual statements imply opinions. Politics is another can of worms. Here, the current aspect extraction algorithms only had limited success because few political issues (aspects) can be described with one or two words. Political sentiments are also harder to determine due to complex mixture of factual reporting and subjective opinions, and heavy use of sarcastic sentences.”

A closer analysis of opinionated texts (See Chapter 3 on Corpus of Opinionated Texts) shows that the argumentation structure does not always resemble the standard forms of logical reasoning which states that every argument comprises at least one intended thesis statement/claim and one or more supporting statements, called premises (Dowden, 1993). Neither does the argumentation structure in such texts necessarily fit into one of the argumentation schemes as discussed by (Toulmin, 1958). Toulmin's argumentation scheme can be described as – Given the data, which are claims or facts or opinions and provided the warrant or justification of the data as

evidence, we can draw a conclusion about the truth of the data. Alternatively, given the data and provided the warrant, unless a rebuttal or doubt or counterargument is put forward against the data, we can draw a conclusion about the truth of the data.

To make things even more complicated, the possible association of the strength of an argument with the standard notion of validity also does not exactly fit in case of such texts, which fall under persuasive texts. A valid argument from the argumentation perspective cannot be assumed a strong one in such texts. There are other parameters like specific exceptions and exclusions that would need to be taken into considerations as argued by (Kolflaath, 2007). Such exceptions and exclusions serve as cases raising conditions for the claims to be not true.

In persuasive texts, in order to perform a discourse analysis, one needs to take into consideration several correlated aspects like text, author, purpose, audience and context as opined by (Burton, 2014) . In his own words, “a speaker or writer takes into account the contingencies of a given place and time, and considers the opportunities within this specific context for words to be effective and appropriate to that moment.” This encompasses rhetorical analysis, an important domain in text analysis. Furthermore, under discourse analysis, one needs to deal with important concepts like discourse constituent units (DCU), discourse segments etc. as discussed in (Polyani, 2003).

From the above discussion, it follows that the discourse analysis and in particular synthesis of opinions and arguments across opinionated texts is an interesting and a motivating research topic. The major challenges of such a task lie in the following:

- 1) Identification of the argumentation structure and the strength exhibited by the arguments.
- 2) Identification of the discourse units and the rhetorical relations within the arguments.
- 3) Identification of the relatedness of any argument with respect to a global thesis statement or issue of discussion.

- 4) Identification of the type of support (“For” or “Against”) that any support statement exhibits with respect to a thesis statement.
- 5) Organization of the supports across multiple opinionated texts over a common topic with respect to a thesis statement or issue of discussion.

This dissertation brings together the relevant concepts from closely related sub-disciplines of discourse analysis like Argumentation Theory, Logic and Practical Reasoning, Rhetorical and Discourse Analysis, Persuasion Theory and Persuasiveness, Information Retrieval, Opinion Mining, Opinion Summarization and Opinion Search and Retrieval for coming up with partial solutions to the given task.

1.2. Objectives and goals

The overall objective of this work is to create methods and algorithms to consolidate material from a number of editorials on a particular topic, taking into account the opinion orientations of the views expressed in these editorials. Arising from this objective, there are a number of sub-objectives, which are:

- a) To analyze the structure of real editorials as found in news-media
- b) To determine methods for analyzing the opinions embedded in editorials in both the lower levels (word and phrase) and higher levels (sentence and paragraph).

In line with the objectives and goals, this thesis intends to answer the following research questions:

- i. What constitutes the argumentation structure in editorials? Is it generic or varies from one editorial to another?
- ii. What are the issues and challenges in analyzing opinions and arguments in editorials?
- iii. What does it involve to carry out a synthesis of opinions and arguments from a single or multiple editorials?

1.3. Research method

To find out answers to my research questions in the previous section, I take the following approach:

- I. By surveying the literature, I come up with an initial method of analyzing opinions and arguments in opinionated texts.
- II. I, then manually analyze a large number of typically opinionated texts, in this case, newspaper editorials using this initial method.
- III. Based on the experience from manual analysis and the knowledge gathered from the available literature, I develop a more refined analysis method, namely the semantic tagset and a set of tagging guidelines, which consists of a more or less exhaustive set of attributes and instructions to capture opinions and arguments in opinionated texts.
- IV. I test the validity as well as coverage of the semantic tagset by involving human annotators for manually annotating real texts from the opinionated texts domain. Through the inter-annotator agreement study, I confirm that the task can be performed with reasonable accuracies.
- V. I then look into the existing linguistic resources and techniques and use them in realizing an automatic tagger application software using the semantic tagset for annotating opinions and arguments in opinionated texts.
- VI. In order to address the shortcomings or inadequate knowledge in the existing linguistic resources, I augment the former with my own set of domain-specific knowledge and again reevaluate the quality of the annotations compared to the manual annotations made by human annotators.
- VII. I repeat the steps from III-VI until satisfied.
- VIII. Finally, I report the overall evaluation of the performance of the tagger application with respect to the different tagging attributes of the semantic tagset.

- IX. Next, for the task of developing a synthesis of opinions and arguments from single or multiple sources, I develop a method to extract the related texts\paragraphs from texts with respect to the thesis statement. This method involves developing query expressions based on the terms of the thesis statement. Again, with the theoretical knowledge gathered from literature on information retrieval as well as practical experiences dealing with text retrieval, I come up with the idea that the query expression can be composed of three categories of terms – exact term, related term and/or domain-specific term to the terms of the thesis statement. Through experiments on the texts, I come up with a method of deriving related and domain-specific terms for the terms of any given thesis statement with examples.
- X. I, then, evaluate the accuracies of the retrieved texts in terms of relatedness with the thesis statement thereby in turn evaluating the effectiveness of the query formulation process. I validate that involving one or more combination of related and domain-specific terms can yield better accuracies in terms of retrieving related texts with respect to any given thesis statement. I rank the retrieved texts for relatedness by the overall relatedness score measure, which is the combined measure provided by the product of the relatedness scores of the individual query terms found in the texts under consideration and their occurrence in the documents with respect to the total number of documents in the corpus.
- XI. Once I retrieve related texts, I exploit the useful information associated with the texts like source, title, date and text/paragraph number to further look for these text fragments in annotated texts.
- XII. After locating the concerned text/paragraph in the annotated texts, I extract the annotations for opinion and arguments together with the text fragments and organize them together in a synthesized textual representation form.

1.4. Outline of the dissertation

The dissertation comprises seven chapters. In this chapter, I present the motivations of the study, the objectives and the goals. I also discuss the research method adopted for the given work in this chapter.

In the second chapter, I provide a literature review on topics and problem domains relevant to my task of opinion and argument analysis/synthesis in opinionated texts. I start with a brief overview of the problems and challenges of the given task in the light of current research in the Opinion Mining problem domain. Then I define the structure of opinionated texts and persuasive writing, consulting a number of resources available in literature. I also discuss the nature of persuasive texts. Then I look into the different prevailing argumentation models and theories thereby trying to define the scopes and composition of arguments in opinionated texts. I conclude the chapter with a discussion on discourse analysis and how it can be applied to analyzing opinionated texts.

In the third chapter, I discuss the argumentation structure of opinionated texts taking examples of real texts from the opinionated corpus and the challenges from the perspective of opinion and argumentation analysis. As a means to develop annotated resources for analyzing opinions and arguments in opinionated texts, I propose a semantic tagset to annotate the texts from the corpus. I also present an inter-annotator study that illustrates the general level of difficulty of the annotation work.

In the fourth chapter, I look into the different linguistic resources (the existing ones and that I have developed or compiled) and their overall contribution in the different tasks of opinion and argument analysis, viz., subjectivity detection, determining the opinion orientation, opinion strength, commitment or intent level of statements, rhetorical relation between statements etc. I also discuss the evaluation of the results in terms of use of the different linguistic resources for the different tasks mentioned above.

In the fifth chapter, I discuss the overall task of constructing a synthesis of opinions and arguments from multiple sources with respect to a common issue or topic of discussion. I look into the three sub tasks, namely, the retrieval of related texts with respect to the thesis statement or the issue of discussion, the extraction of opinion and argument information for the related text fragments from the annotated texts and finally the construction of the synthesis of information out of the extracted information.

In the sixth chapter, I present the computational linguistic model or a proof of concept for the analysis and synthesis of opinions and arguments from opinionated texts. The chapter discusses the different components of the model and connects the computational and linguistic aspects of the given task together.

Finally, in the seventh chapter, I summarize the contributions of this work and present the current limitations and recommendations for future works.

CHAPTER 2 LITERATURE REVIEW

As presented in Chapter 1, Opinion Analysis/Synthesis in opinionated texts encompasses a number of disciplines namely, Argumentation Theory, Logic and Practical Reasoning, Rhetorical and Discourse Analysis, Persuasion Theory and Persuasiveness, Information Retrieval, Opinion Mining, Opinion Summarization and Opinion Search and Retrieval. This Chapter discusses the different aspects from these disciplines related to the problem thus drawing upon knowledge to seek partial solutions. First of all, Opinion Mining and Related Problems are discussed, then after gradually moving into Editorials and Persuasive Writing, Argumentation Theory and Discourse and Rhetorical Analysis.

2.1. Opinion Mining and Related Problems

Although, Opinion Mining/Sentiment Analysis has emerged only quite recently as a sub-discipline under Computational Linguistics, a considerable work has been done in this direction. These works range from a variety of task domains like mining the product reviews available on the web, sentiment classification of documents, opinion mining and summarization, opinion retrieval etc. A detailed survey of the works on Sentiment Analysis including the most recent ones is presented in (Liu, 2012). The works can be classified as dealing the problems in three broad levels – the document level, the sentence level and the clause level. In the document level, an opinion document is classified as expressing a positive or negative opinion or neutral sentiment. Similarly, in the sentence level, a sentence is first classified as expressing an opinion or fact and then after whether the opinion expressed is positive, negative or neutral. Finally, in the clause or phrase level, opinions expressed with respect to a particular target object and their polarity or orientation and strength are studied. Irrespective of the nature of the different specific tasks, Opinion Mining generally encompasses the following generic problems:

- I. Determining the subjectivity or identifying the subjective and objective expressions in texts where subjective expressions are generally opinions whereas the objective

ones denote facts. Given below are two examples of subjective and objective sentences(Liu, p.27):

Subjective sentence – “I like iPhone.”

Objective sentence –“iPhone is an Apple product.”

Further, Liu in his book makes a distinction between “subjectivity” and “sentiment”, where he argues that these two terms are not really the same. Sometimes a subjective sentence may not express any sentiment at all as in “I think that he went home” whereas at other times even an objective sentence can imply opinions or sentiments due to desirable and undesirable facts (Zhang & Liu, 2011) as in “The earphone broke in two days.” This makes determining the subjectivity a difficult task. The task is also known as subjectivity classification in literature.

Wiebe, Bruce, & O'Hara (1999) and Bruce & Wiebe, (2000) manually annotated a corpus of 1,001 sentences of the Wall Street Journal Treebank Corpus (Marcus, Marcinkiewicz, & Santorini, 1993). Three human judges were involved to label each sentence as subjective or objective. The Expectation-Maximization (EM) algorithm was used to produce the correct tags based on the consensus opinions of the judges. Out of the total 1,001 sentences, 486 were subjective sentences whereas 515 were objective sentences. The work of Bruce & Wiebe showed that adjectives and subjective sentences have a statistically significant and positive correlation; the probability of a sentence being subjective based on the presence of at least one adjective was found to be 55.8%.

Wiebe, Bruce & O'Hara perform subjectivity classification using the Naïve Bayes classifier in which the authors distinguish sentences used to present factual information from sentences used to present opinions and evaluations. The authors use a set of binary features, e.g., the presence of a pronoun, adjective, cardinal number, modal other than “will” and adverb other than “not” to train their classifier. They report the accuracy results of their classifier as being 72.17% as compared to 81.5% that by human annotators.

Wiebe (2000) uses an unsupervised method for subjectivity classification in which an initial set of seed words denoting subjectivity are expanded with similar words indicating subjectivity. The author also uses gradable adjectives, a concept first proposed by (Hatzivassiloglou & Wiebe, 2000) for determining subjectivity. Hatzivassiloglou & Wiebe (2000) establish that dynamic adjectives, semantically oriented adjectives and gradable adjectives are strong predictors of subjectivity . The authors differentiate words having respectively positive and negative orientation by looking at whether they encode a desirable state (e.g. beautiful) versus an undesirable state (e.g. ugly).

- II. Determining the orientation or polarity of the subjective expressions. Hatzivassiloglou & McKeown (1997) exploit linguistic rules on connectives to determine the orientations of sentiment words from a corpus. While the conjunction AND marks two adjectives of the same orientation as in "This car is beautiful and spacious", a BUT connective indicates adjectives of different orientations as in "This car is beautiful but difficult to drive". Similar other rules were developed for other connectives OR, EITHER-OR, NEITHER-NOR. A graph linked adjectives of same and different orientations, after which clustering was done to produce positive and negative adjectives.

Pang, Lee, & Vaithyanathan, (2002) use supervised learning methods to classify movie reviews into positive and negative classes. The unigrams (a bag of words) features were found to be quite useful with both naive Bayes or Support Vector Machines (SVM) in terms of the classification.

Turney (2002) classifies opinions based on some fixed syntactic patterns that are likely to be used to express opinions. These patterns constitute certain sequence of part-of-speech (POS) tags. The approach follows a two step algorithm:

Step 1: Extract two consecutive words if their POS tags match to any of the patterns presented in Table 2-1

Table 2-1: Patterns of POS tags for extracting two-word phrases

	First word	Second word	Third word (not extracted)
1	JJ	NN or NNS	Anything
2	RB, RBR, or RBS	JJ	not NN nor NNS
3	JJ	JJ	not NN nor NNS
4	NN or NNS	JJ	not NN nor NNS
5	RB, RBR, or RBS	VB, VBD, VBN, or VBG	Anything

Source: (Liu, Sentiment Analysis and Opinion Mining, 2012)

Pattern 1 in the above table allows the extraction of two word sequences, whereby the first word is an adjective, the second is a noun and the third, which will not be extracted, could be anything. In the example, “The mountains offer a majestic scene”, “majestic scene” is extracted as per pattern 1. It is noted that in the patterns above, adjectives (JJ) and adverbs (RB, RBR and RBS) characteristically express opinions.

Step 2: For the extracted phrases, the sentiment orientation (SO) is estimated using the pointwise mutual information (PMI) measure:

$$PMI(term_1, term_2) = \log_2 \left(\frac{Pr(term_1 \text{ and } term_2)}{Pr(term_1)Pr(term_2)} \right) \quad (2.1)$$

PMI gives a measure of the degree of statistical dependence between two terms. Here, $Pr(term_1 \text{ and } term_2)$ is the actual co-occurrence probability of $term_1$ and $term_2$, and $Pr(term_1)Pr(term_2)$ is the co-occurrence probability of the two terms when they are statistically independent. Once the PMI of two terms is

calculated, the sentiment orientation (SO) of a phrase is computed based on its association with the words “excellent” and “poor”, which have respectively a positive and a negative orientation.:

$$SO(phrase) = PMI(phrase, "excellent") - PMI(phrase, "poor") \dots \dots \dots (2-.2)$$

For calculating the probabilities, queries are issued to a search engine and the number of hits is collected, where hits refer to the number of relevant documents to the query. In this case, the AltaVista search engine was used as it has a NEAR operator, which constrains the search to documents containing the words within ten words of one another in any order.

Let $hits(query)$ be the number of hits returned, then Equation (2.2) can be rewritten as:

$$SO(phrase) = \log_2 \left(\frac{hits(phrase \text{ NEAR } "excellent")hits("poor")}{hits(phrase \text{ NEAR } "poor")hits("excellent")} \right) \quad (2.3)$$

Step 3: Given a review document, the average SO of all phrases in the review are computed. The review is then classified as positive or negative according to whether the average SO is positive or negative.

It has been reported that the final classification accuracies have been 84% for automobile reviews and 66% for movie reviews.

Esuli & Sebastiani (2005) use supervised learning for the classification of words into positive and negative classes. The authors initially collect a set P of positive seed words and a set N of negative seed words. Then they expand the sets using synonym and antonym relations in an online dictionary (e.g., WordNet) thus generating the expanded sets P' and N' , which form the training set. The glosses in

the dictionary for each term in $P'UN'$ are then used to generate a feature vector. Based on the feature vector, a binary classifier is then built using different learning algorithms. A third category “objective” was also included and the objective seed set was expanded using hyponyms in addition to synonyms and antonyms. Different strategies were then adopted to do the three-class classification. Esuli & Sebastiani (2006) build the SentiWordNet following the same method described above. For each synset of WordNet, the SentiWordNet is associated with three numerical scores $Obj(s)$, $Pos(s)$ and $Neg(s)$ that respectively describe the degree of objectivity, positivity and negativity of the synset terms. Similarly, Kim & Hovy, (2006) initially take three seed sets of positive, negative and neutral words. The set is then expanded with synonyms from WordNet. Finally, they use a Bayesian formula to determine to which class a particular word belongs to.

- III. Determining the strength or intensity of the subjective expressions. Wilson, Wiebe, & Hwa (2006) use machine learning techniques to train classifiers to predict the intensity of clauses and sentences. The learning algorithm uses a large lexicon of subjectivity clues², which are words and phrases that may be used to express opinions and emotions. The authors organize the clues in the lexicon into sets and create one feature per set. The clues are used to define the intensity of the clauses and sentences for the experiments. They perform 10-fold cross-validation experiments using three different learning algorithms: boosting, rule learning, and support vector regression. The authors show that many clues of the subjective language, including syntactic clues can be used for the task of intensity recognition. The best results for intensity classification are achieved when the widest variety of the clues are used.

The problem of opinionated texts analysis involves all the problems I-III above.

² http://miqa.cs.pitt.edu/lexicons/subj_lexicon/

Next, we move on to two important areas of research in Opinion Mining, namely Opinion Summarization and Opinion Search and Retrieval, both of which are related to the problem of opinionated texts analysis/synthesis. Opinion Summarization carries some similarity with traditional text summarization in that it also deals with multiple text documents. However, in the case of Opinion Summarization, the summary is more focused on entities and aspects and the sentiments being expressed over them (Liu, 2012). In addition to this, the opinion summary is also associated with a quantitative aspect about the opinions, i.e., the number of percent of people holding positive or negative opinions about the entities and aspects. Hu & Liu (2004) generate feature-based summaries of customer reviews of products sold online. Given a set of customer reviews of a particular product, the authors i) identify the features of the products that customers have expressed their opinions on, ii) for each feature, identify the review sentences that give positive or negative opinions, and finally, iii) produce a summary using the discovered information.

Below in Figure 2-1, an example summary of a particular camera, Digital_camera_1 is provided:

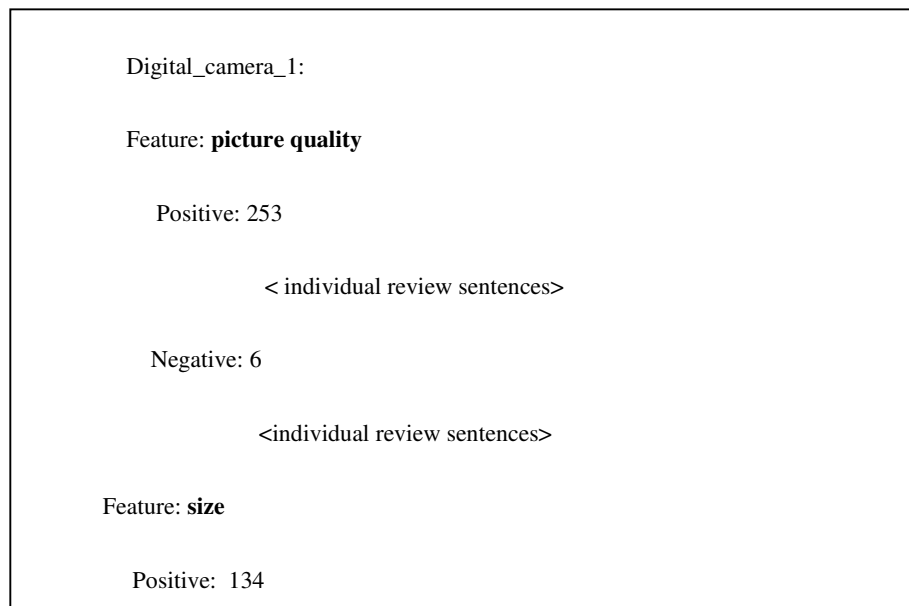


Figure 2-1: An Example Summary

In Figure 2-1 above, the summary is in a structured text form where the different opinions expressed by the reviewers on the particular product, “Digital_camera_1” have been summarized according to the different features of the product. The summary provides not only the total positive and negative counts in the reviews but also an account of the individual review sentences under each of the positive and negative categories. For task 1, i.e., mining product features, the authors use both data mining and natural language processing tasks. For task 2, i.e., identifying opinion sentences in each review and determining the opinion orientation, they rely on a list of adjectives together with WordNet. For task 3, i.e., summarizing the results, the authors put the sentences into positive and negative according to the opinion sentence’s orientation. A count is computed to show the total number of reviews giving positive/negative opinions to the feature. Next, the features are ranked as per their frequencies of occurrences in the reviews.

Opinion Search can be broadly categorized into two types (Liu, 2012):

1. Finding public opinions about a particular entity or aspect of entity. For example, finding customer opinions about a product or the feature of a product, and find public opinions about a political issue or candidate.
2. Finding the opinions of a person or organization (i.e., opinion holder) about a particular entity or an aspect of the entity (or topic). For example, NATO’s view about Russia’s involvements on the political developments in Crimea.

Opinion search shares with traditional web search the following tasks:

1. Retrieving relevant documents/sentences to the user query.
2. Ranking the retrieved documents or sentences.

Additionally, opinion search needs to determine whether the documents or sentences express opinions on the query topic (entity and/or aspect) and whether the opinions are positive or negative. Opinion retrieval is currently undertaken in two stages. The first stage ranks the documents by topic relevance only. Next, in the second stage, the candidate relevant documents from the first stage are re-ranked based on their opinion scores. However, the two stages are

combined in one in more advanced models whereby the rankings are based on the integrated scores of the topic relevance and opinion at the same time.

In their Opinion search or retrieval system, Zhang & Yu (2007) work with two components, namely, the retrieval component and the opinion classification component. The retrieval component performs the traditional information retrieval (IR) task considering both keywords and concepts. Concepts refer to named entities(e.g., names of people or organization) or various types of phrases from dictionaries and other sources. Hence, the first step would be to identify the concepts within the user query and then the search query terms are expanded with its synonyms. Further, concepts in the retrieved documents are also identified and relevant words are automatically extracted from the top-ranked documents. Finally, for each retrieved document, a similarity or relevance score is computed with respect to the expanded query. The opinion classification component, on the other hand, performs two tasks – i) classification of the document into one of the two categories, opinionated and not-opinionated, and ii) classification of each opinionated document as expressing a positive, negative, or mixed opinion. Both the tasks use supervised learning. A large amount of opinionated (subjective) training data is taken from review sites and other domains like consumer goods and services. Similarly, the non-opinionated training data are taken from sources like the Wikipedia. From the above training data, a SVM classifier is constructed, which performs the above mentioned two classification tasks. The opinionated sentences in a document and their strengths together with the document's similarity with the query are taken into consideration while ranking the document.

The study and analysis of the works on Sentiment Analysis in this section reveals that much of the efforts are still too focused in the lower levels of opinion mining, i.e., in the lexical, phrase and sentence levels. Such lower levels of analysis alone would not be sufficient for the problem of Opinion and Argument Analysis and Synthesis in opinionated texts as it necessarily requires the analysis of opinions and arguments in much higher level, i.e., in one to several paragraphs (discourse level). In the next section, I discuss Editorials and Persuasive Writing, where I highlight the structure of editorials and persuasive writing and make a first step towards looking into

problems of opinion and argumentation analysis in opinionated texts at the paragraph level and above.

2.2. Editorials and Persuasive Writing

Editorials, which represent opinion articles written by publishers, editor(s) or columnist(s) of newspapers, 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. The language in such persuasive texts is influenced by at least the following factors (Roberts-Miller, 2008; Government of South Australia, 2011):

- Purpose
- Audience
- Identity
- Attitude

Since the primary purpose of persuasive texts is to persuade the audience over a particular stand or position or viewpoint, the contents of the texts varies depending upon the intentions of the author(s) and audiences the author is targeting. In the list of factors above, identity refers to the role of the author that he/she projects himself/herself as to the audience. In the similar manner, attitude denotes the kind of emotions (unhappiness, concern, worry etc.) that are expected to be generated on reading the persuasive text.

Roberts-Miller discusses that persuasive writing or rhetoric is guided by three terms, which can often be depicted in the form of a triangle – **text**, **author** and **audience**. She further reports that the **triangle** is often associated with two more terms – **the immediate context** and **the larger context** (also known as **background**). Miller further argues that the text is influenced by the intention of the author as well as to what he/she thinks that audience knows about the topic. Furthermore, it is also dependent upon the **contexts** in which the writing was developed.

Persuasion and persuasion theory is a well established research area in the field of social sciences, mass media and communication theory, psychology and related disciplines. Simons (1976, p.21) defines persuasion as “human communication that is designed to influence others by modifying their beliefs, values, or attitudes”. O’Keefe (2002) argues that persuasion involves at least three things – the sender or message conveyer, the means for conveying this message and the recipient of the message. On part of the message sender, he or she has clearly a goal which needs to be achieved. Communication is the means to achieve the goal and finally the message recipient must take the message in free will rather than under the influence of force. Persuasion in general is closely associated with the term “attitude”. Simons, (1976, p.80) defines the term “attitude” as a “relatively enduring predisposition to respond favorably or unfavorably” towards something. O’Keefe (2002) asserts that each one of us have attitude towards people, places, events, products, policies, ideas etc. Dainton & Zelly (2013) argue that attitude are learned evaluations and not something that people are born with, thus they are subject to change. Furthermore, they are presumed to influence behavior. Some of the well known persuasion theories are as follows – i) cognitive dissonance, ii) social judgement theory, iii) inoculation theory, iv) balance theory, v) rank’s persuasion model, vi) source credibility theory, vii) congruity theory, vii) belief congruency, viii) reinforcement theory, ix) information manipulation theory, x) elaboration likelihood and xi) attribution theory.

2.3. Argumentation Theory and Structure

Now we will look into theories and techniques for analyzing and extracting the texts/paragraphs from the editorial texts. Of central focus among these theories and techniques are Argumentation Theory and Argumentation Structure in Argumentative texts.

Argumentation Theory has roots from a number of interdisciplinary fields like philosophy, logic, linguistics, law, Artificial Intelligence etc. In general, three major models (Toulmin, 1958; Van Emmeran & Grootendorst, 2004 and Walton, Reed, & Macagno, 2008) are found to dominate the field. The last two models relate to dilectical argumentation and hence are not relevant to the given thesis topic as opinionated texts are necessarily monological argumentation. Besnard &

Hunter (2008) make a distinction between monological and dialectical argumentation systems in that "dialectical argumentation is conducted between two or more agents and involves more complex issues of multi-agent interaction, dialogue protocols and strategies" whereas in case of monological argumentation, such a multi-agent interaction and dialogues are not necessarily present, rather the views of a single entity are evident in the form of text or other medium. Toulmin's model is briefly described below.

Toulmin defines the layout of arguments and argues that an argument comprises at least three components – data, warrant and claim. Data refers to the facts or opinions on evidences whereas claim refers to the conclusion reached via the warrants, which are suitable justifications or explanations based on the data and leading to the claim. In Toulmin's words, 'an argument is a movement from accepted data through a warrant to a claim'. He further adds three other components to the layout – backing, rebuttal and qualifier, which he terms as not mandatory but may be present. Backing refers to the evidence or support for the assumptions made in the warrant. Rebuttal refers to the conditions which would lead the claim to be false whereas the qualifier denotes the probability of the confidence level of the claim being true. Toulmin's argumentation model can be represented in abstract syntax using extended Backus-Naur Form or EBNF in the form of production rules as shown in Figure 2-2 below:

Toulmin-argument \rightarrow Data, Warrant+, Claim;

Data \rightarrow Proposition;

Warrant \rightarrow Predicate, Backing;

Backing \rightarrow Predicate;

Claim \rightarrow Qualifier, Proposition, Rebuttal;

Qualifier \rightarrow Predicate;

Rebuttal \rightarrow Predicate;

Proposition \rightarrow (* a simple expression that is claimed to be true, such as “X is a citizen of Bermuda”*)

Predicate \rightarrow (* a more complex logical expression, such as “if X is a citizen of Bermuda, then X is a British Subject” *)

Figure 2-2: Abstract Syntax Representation of Toulmin’s Model

Toulmin’s model of argumentation and the discussion on the composition of an argument is close to the two standard forms of propositional logic popularly known as Modus Ponens and Modus Tollens, respectively: i) Modus Ponens (latin for “mode that affirms”), which states as “Given the Major and Minor Premises (Hypotheses) are True, the Conclusion is True” and ii) Modus Tollens (latin for “mode that denies”), which states as “Given a Negative Minor Premise or Hypothesis and a Major Hypothesis or Premise are True, a negative Conclusion about the antecedent or the Major Premise or Hypothesis is True.” Examples to each of the two forms of logic are provided below:

Modus Ponens – Example (Rosen, 1998, p.63)

Major Premise: “If you have a current password, then you can log on to the network.”

Minor Premise: “You have a current password.”

Therefore:

Conclusion: “You can log on to the network.”

Modus Tollens – Example (Rosen, 1998, p.63)

Major Premise: “If you have a current password, then you can log into the network.”

Minor Premise: “You can’t log on to the network.”

Therefore:

Conclusion: “You don’t have a current password.”

From the perspectives of argumentation, opinionated texts are organized in single text files talking on a particular issue or topic of discussion. These single text files comprise one to several paragraphs, each of them representing a single composite argument or idea. In this respect, each paragraph has an opening statement or thesis analogous to the Claim in Toulmin’s model of argumentation followed by one or more Data and Warrants (supporting statements or simply supports). Within the paragraph, there also may be one or more counter-arguments with respect to the Claim. However, since the argumentation approach in opinionated texts is not necessarily dialectical (discussion involving more than one party) but monological, such counter-statements possibly raising doubts about the supports being put forward in favor of the claim is not a common phenomenon. In Figure 2-3, the text structure of editorials is presented in the form of abstract syntax.

Editorial \rightarrow Date, Source, Topic, Gist?, Paragraph+

Paragraph \rightarrow Opening-Statement-thesis, Support+

Opening-Statement-thesis \rightarrow Statement-Paragraph

Support \rightarrow Statement-Paragraph+

Statement-Paragraph \rightarrow Sentence

Sentence \rightarrow Phrase+

Phrase \rightarrow Word+

Figure 2-3: Structure of editorials in Abstract Syntax Representation

As shown in Figure 2-3 above, an Editorial comprises Date of publication, Source of publication or name of the newspaper, a Topic, optionally (zero or one) a Gist which is generally a one liner or summary of the editorial and one or more Paragraphs. A Paragraph comprises the Opening statement-thesis and one or more Support. The Opening statement-thesis is a Statement-Paragraph and the Support represents one or more Statement-Paragraph. A Statement-Paragraph is a Sentence and consists of one or more Phrase. Finally, a Phrase comprises of one or more Word.

With the structure of editorials defined above, now we will look into Discourse and Rhetorical Structure Theory in the next section as a means of analyzing and synthesizing the opinions and arguments.

2.4. Discourse Analysis

As mentioned in the earlier sections, sentences or statements within the same or different paragraphs in the same or different opinionated texts may be logically connected to each other. These connected or coherent statements form part of the argument(s) with respect to the opening statement or claim and hence represent the discourse. From this, it follows that in order to correctly analyze the arguments and consequently develop a synthesis of the arguments, it is essential to detect and determine the rhetorical relations in the opinionated texts. The rhetorical relations are also interchangeably referred to as discourse or coherence relations in literature. Rhetorical Structure Theory (RST) is one of the most widely used theories for describing the model of text organization and deals with these relations (Mann & Thompson, 1987). RST defines a set of 23 rhetorical relations, which can exist between spans of texts within a discourse. The central focus of RST is on two broad categories of text spans, namely, a **nucleus** and a **satellite**. While the nucleus represents the writer's purpose and hence is something that can be interpreted independently, the satellite is not independent and can be interpreted only with respect to the nucleus. For my task of opinionated texts analysis and synthesis, as a result of manual analysis of the corpus of opinionated texts (See Chapter 3 for details of the corpus study), I have concluded that out of the twenty-three rhetorical relations originally proposed by Mann & Thompson, the following 8 rhetorical relations in general suffice for establishing the relationship between the opening statement-thesis and the supports as well as between supports – i) Elaboration, ii) Result, iii) Reinforcement, iv) Contrast, v) Cause-Effect, vi) Exemplification, vii) Conditional, viii) Explanation. More details on these eight rhetorical relations will follow in Chapter 3.

Now let us move on to discuss some of the works in the direction of automatic discourse segmentation and text coherence, both of which are crucial for our task. For discourse segmentation, which involves segmentation of paragraphs and coherent clauses or sentences, the role of discourse markers or cue words is found to be very vital. Essentially these markers or cues represent word(s) or phrases that function to signal discourse structure.

There has been adequate study of the discourse markers and their role in discourse analysis. Taboada (2009) report that coherence relations in texts are often signalled via cue phrases called discourse markers or particles. By discourse markers, the author refers to any conjunction, adverb, adverbial phrase that frequently links two or more units of discourse. However, equally possible are situations where these markers do not explicitly occur but must be understood by context or implicitly. In addition to this, the discourse markers are often ambiguous in terms of conveying the rhetorical relations involved. For example, the discourse marker “Since” can be related to both “Temporal” and “Explanation” rhetorical relations depending upon the contexts of their occurrences (Sporleder & Lascarides, 2008). Such situations amplify the complexity of automatically detecting discourse relations in texts. Besides discourse markers, the author also notes that there are other signalling mechanisms like – i) Morphological (verb tenses indicating temporal relations), ii) Syntactic (sentence mood – imperative, indicative, interrogative), iii) Semantic (causal relations indicated by verbs like “cause”, “trigger”, “provoke” etc.), iv) Pragmatic (implicature relations implying certain meanings even in the absence of explicit cues). Fraser (1999) present a number of canonical forms of the patterns of Discourse markers (below, DM stands for Discourse Markers and S1 and S2, respectively represent sentence 1 or clause 1 and sentence 2 or clause 2:

<S1. DM + S2>

For example:

The people are happy with the Prime Minister. **However**, the opposition leaders are not.

<S1, DM+S2>

For example:

I read a book, **and** he played football.

<DM+S2, S1>

For example:

Since this is first journey by airplane, you may feel nervous.

Power, Doran, & Scott (2000) note some of the syntactic and rhetorical properties of discourse markers:

“However” can occur in several positions.

“Since” and “Consequently” can appear anywhere.

Discourse markers, hence, can serve as effective sign-posts to signal the presence of discourse or coherence or rhetorical relations in discourse. From the study of discourse markers in the literature (Sharpling, 2012; London Metropolitan University, 2010; J.J. Strossmayer University of Osijek, 2008; Sweneey, 2010), I have ended up with a total of 19 different relations, namely, Addition, Cause-Effect, Comparison, Concession, Conclusion, Contrast, Repetition, Sequencing, Similarity, Structure, Temporal, Counter-argument, Emphasis, Exemplification-Illustration, Express-attitude, Generalization, Logical-consequence, Reinforcement and Explanation. Similarly, for each of these relations, I have compiled different connectors or connectives of varying word counts and lengths ranging from unigrams (“also”, “as” etc.) up to five grams (“as a matter of fact”). So that the identified relations are in conformance with the previously established set of eight rhetorical relations, I have done the remapping of the above relations in the following forms:

{Temporal, Sequencing, Structure, Comparison, Similarity} → {Elaboration}

{Logical-consequence, Conclusion, Generalization} → {Result}

{Repetition, Emphasis, Express-Attitude, Addition} → {Reinforcement}

{Contrast, Counter-argument} → {Contrast}

{Cause-Effect} → {Cause-Effect}

{Exemplification-illustration} → {Exemplification}

{Condition, Concession} → {Conditional}

{Explanation} → {Explanation}

Here, I have considered only the explicit discourse markers or connectives in order to identify or categorize the different discourse/rhetorical/coherence relations in the analysis/synthesis of opinionated texts. As mentioned earlier, the explicit discourse markers or connectives occur explicitly in texts and represent the different rhetorical relations presented above.

2.5. Summary

In this Chapter, a literature review was conducted on a number of topics and problem domains relevant to the task of opinion and argument analysis/synthesis of opinionated texts. I started with a brief overview of the problems and challenges of my task in light of the current research in the Opinion Mining problem domain. It should be noted that much of the works in Opinion Mining are still too focused on the lower levels, i.e., word, phrase and the sentence levels. However, to conduct the analysis and synthesis of opinionated texts, we would need to move to higher levels, i.e., paragraph and higher levels. This merits the discussion of the structure of editorials and persuasive writing. Hence, after the discussion on Opinion Mining problems, I moved on to defining the structure of opinionated texts and particularly editorials and persuasive writing. The discussion unfolded that the structure of opinionated texts comprises one to a few paragraphs with each paragraph consisting of the opening statement or thesis and one to a few supporting statements or supports. The next section on Argumentation Theory and Structure in Opinionated Texts discussed the Toulmin argumentation model thereby defining the scope and putting into context the composition of arguments in opinionated texts. Based on this, the different sections of an editorial were outlined and represented in abstract syntax or production rules.

Finally, we looked into Discourse Analysis where the discussion was focused on how Discourse Markers can act as useful devices to establish discourse or coherence or rhetorical relations in discourse.

This Chapter looked into the problem of analyzing and synthesizing opinions/arguments by going through the literature available on different closely related disciplines. We learnt that the problem in itself is a very challenging and not a straightforward one – to come up with some partial solutions; one needs to borrow concepts from different disciplines. However, we are yet to verify the knowledge that we have gathered about opinionated texts in real textual settings – whether they exhibit the same nature, pose the same set of challenges or require some different treatment for processing them in terms of analysis or synthesis. The next Chapter will look into the Corpus of Opinionated Texts thus verifying whatever we have gathered from literature about Opinionated Texts and at the same time will bring forth other issues pertinent to the analysis of such texts in real textual settings. The Chapter will also work towards building a means for creating a linguistic basis for analyzing such texts, namely, developing a semantic tagset for annotating the texts, that will serve as a useful resource for analysis and synthesis.

CHAPTER 3 CONSTRUCTING THE CORPUS

In Chapter 2, we looked into the problem of analyzing and synthesizing opinions/arguments by going through the available literature on different closely related disciplines. We saw that the problem in itself is a very challenging and not a straightforward one requiring us to borrow concepts from different disciplines in order to find partial solutions. However, we are yet to verify the knowledge that we have gathered on opinionated texts via the literature review in real textual settings. In this Chapter, we will look into the construction of the corpus thus verifying whether the knowledge we have gathered from literature holds true about Opinionated Texts and at the same time, we will bring forth other issues pertinent to the analysis of such texts in real textual settings. We will also work towards building a means for creating a linguistic basis for analyzing such texts, namely, developing a semantic tagset for annotating the texts, that will serve as a useful resource for analysis and synthesis.

3.1. Motivations behind building a Corpus

From the first two chapters, we know that analyzing the opinions and arguments in opinionated texts is not a very straightforward task. This leads us to investigate further the following questions on the opinionated texts in real textual settings:

- i) What is the argumentation structure of opinionated texts? Does it comply with the structure as prescribed in the literature?
- ii) What does it involve to conduct a discourse analysis of a single or multiple opinionated texts?

To arrive to possible answers to the above questions, I build a corpus of opinionated texts. The corpus includes opinionated texts from different sources and belonging to a certain time interval. The corpus not only gives us an idea about the prevailing argumentation structure in opinionated texts but also serves as the source to know the stances or view points of the press agencies over a particular event or topic at a particular time or given time span. It further facilitates us to investigate the specific or general opinion pattern trends pertaining to a single or multiple

opinionated text sources. Findings on such patterns serve for potentially detecting opinions, their polarity or orientation, their intensity or strength and consequently the elements of the argumentation structure in such texts.

For building the corpus, I collect opinionated texts from different national (in the context of Nepal) and international online news portals published in English during the time span 2007-2012. The texts have been collected from five different sources, namely, “Aljazeera”, “BBC”, “The Guardian”, “The Kathmandu Post”, and “Nepali Times”. The first three sources are international whereas the fourth and the fifth are locally published in Nepal, the language medium being English. While selecting the sources, a more or less balanced approach has been adopted in the choice of the sources as per the region wise location. In this regard, the first source belongs to the Middle East, the second and the third represent the United Kingdom, while the rest are published in Nepal. Samples of opinionated texts (both editorials and blogs) from the corpus are presented in Appendix B. In Table 3-1, the information of the collected opinionated texts from different sources is presented.

Table 3-1: Information of the corpus

S.No.	Source	Time span	No. of opinionated texts	No. of words
1	Aljazeera	2011-2012	1830	2624841
2	BBC	2007-2012	853	277344
3	The Guardian	2007-2012	6191	3042377
4	The Kathmandu Post	2007-2012	1718	457828
5	Nepali Times	2007-2012	3173	112064

Corpus based research is not a new endeavor in Computational Linguistics/Natural Language Processing (See for example, NLTK Text Corpora³), however, to the best of our knowledge, this is the first ever developed corpus especially focused on opinionated texts (editorials and blogs) and covering such diverse sources as well as spanning to a considerable time period of six years. In Table 3-2, I present the top 50 frequent discourse markers along with their frequencies of occurrences in the corpus. These markers serve as useful cues in the detection of supports (for example, “Because”, “Since”, “In light of”, “Whereas”, “Given that”, “For the reason that”, “For” etc.) and claims (for example, “Therefore”, “Consequently”, “Hence”, “So”, “Thus”, “In conclusion”, “Accordingly”, “It follows that”, “As a result” etc.) and at the same time also signal the presence of one or more rhetorical relations useful for conducting a discourse analysis. I present the top 200 frequent discourse markers along with their frequencies of occurrences in the corpus in Appendix D.

³ http://www.nltk.org/nltk_data/

Table 3-2: Top 50 frequent Discourse Markers in the corpus

Discourse Marker	Frequency of occurrences in the corpus		
And	52796	However	4252
For	32029	That is	3935
As	26389	Next	2728
But	21247	Such as	2724
So	10723	Once	2721
Also	8923	Though	2683
Like	6639	Often	2570
First	6527	Again	2538
Last	6118	Second	2434
While	5864	Always	2298
Still	5054	Instead	2144
Because	4875	Indeed	1962
Since	4436	Of course	1888
Before	4336	Until	1871
Yet	4301	Although	1812
		Despite	1792
		As well as	1726

Especially	1676
Recently	1576
Later	1567
In fact	1483

Soon	1479
Certainly	1466
Third	1437
Thus	1330

3.2. Argumentation Structure in Texts from the Corpus

Hundred texts, twenty from each of the five sources, respectively, “Aljazeera”, “BBC”, “The Guardian”, “The Kathmandu Post” and “Nepali Times” were analyzed for the argumentation structure. While the latter two sources, “The Kathmandu Post” and “Nepali Times” have a regular “Editorial” section, the first three sources, “Aljazeera”, “BBC” and “The Guardian” have “Opinion” or “Blog” section. Some key observations as part of the analysis are presented below:

- i) The opinionated texts in general have an overall topic with sub-topics within texts as required although in editorials, sub-topics are not explicitly mentioned.
- ii) The opinionated texts may have one or more authors. This is however, not true for editorials in which case the opinion is considered that of the editorial board and hence no names of author(s) are provided.
- iii) The opinionated texts have a publication date.
- iv) The opinionated texts start with the introduction of the issue in the opening paragraph or the lead paragraph.
- v) Each subsequent paragraph discusses the issue under consideration thus trying to provide support(s) or backing to the issue. In this regard, each paragraph represents a single argument. The argument is sometimes unique in the sense that it does not have direct connection with the rest of the arguments whereas in other cases it is connected with the other arguments.
- vi) Arguments related to each other are clustered or grouped together, i.e., organized in adjacent paragraphs (preceding or following) under some sub-topic. Hence, although arguments appearing in such clusters talk about the overall topic or issue under consideration, they are more focused on the theme of the sub-topic, where they are located.
- vii) The concluding argument is the final paragraph(s).

The observations on the argumentation structure of opinionated texts presented above poses the following challenges from the perspective of opinion and argument analysis:

- i) Opinionated texts are usually a mix of facts and opinions. Generally, from an argumentation perspective, any argument appears more convincing and strong if it is aided by facts as supporting statements rather than just opinions. However, making a distinction between facts and opinions itself is a daunting task. Often facts can be opinions in disguise. For example, “The past year will be remembered for violent movements, especially in the tarai, and a few but serious conflicts in the western tarai. Crime and political murders spread fast and hard in the tarai. The Young Communist League (YCL) also set new records in violence and extortion. The year also saw a height of anarchy, impunity and lawlessness.” (The Kathmandu Post, 2008). Here the underlined portions, although seem to be facts by the presence of the verbs “set” and “see” are opinions.
- ii) Determining the standpoints of supports with respect to the claim is difficult because stance or viewpoint is a different concept than opinion orientation. A statement, for example, can be negative in terms of opinion orientation but at the same time can lend a positive support to some statement with a negative orientation. For example, “Nepal is in crisis because of the stagnating economy, falling farm productivity, a rising population and negative job creation.” (Nepali Times, 2009). Here, the underlined portions although have negative opinion orientation render a positive support to the thesis statement “Nepal is in crisis.”
- iii) Intensity or strength of arguments in opinionated texts is determined by a number of elements like intensifiers, pre-modifiers, modal and reporting verbs. However, the situation becomes more complex when more than one of these occurs in the texts in combination.
- iv) Discourse markers or cues play an important role in establishing the connection between statements within a paragraph (argument) and thus between paragraphs (arguments). From this point of view, discourse markers or cues can be useful in

conducting a discourse analysis of the texts. However, in opinionated texts, we do not always find explicitly mentioned discourse markers or cues. There are cases whereby two statements are connected to each other despite the absence of explicit discourse markers. Such statements are said to be connected via implicit markers. It is obvious that dealing with such statements in terms of establishing discourse relations is quite a difficult task.

In order to come up with partial solutions to the given task and the underlying problems, I propose to attack the problem via annotations or tagging of the texts with opinion and argumentation elements basically in the sentence (statement) and higher levels (paragraphs or arguments). Additionally, I propose to annotate useful information related to such texts like source, date, author, issue number, topic, gist etc. Such annotated resources can be used for accurate detection, classification and synthesis of arguments and views in opinionated texts. In the following section, I discuss the semantic tagset that has been developed for annotating the corpus.

3.3. Devising a semantic tagset

There have been growing efforts in developing annotated resources so that they can be used to train the machine and ultimately aid in the automatic identification, extraction and analysis of opinions, emotions and sentiments in texts. Some of such works on corpus annotation, among others, include (Wilson & Wiebe, 2003;2005) and (Read, Hope, & Carroll, 2007).

In the work of (Wilson & Wiebe, 2003), an annotation scheme has been developed for annotating private states in news and other discourse. The private states include expressions of opinions, beliefs, emotions, sentiment and speculation. The design of the annotation scheme has been focused towards addressing the multiple perspectives in question answering systems and thus is primarily directed towards annotating the different forms and situations whereby private states are expressed explicitly or implicitly through the language that a speaker or writer uses and using expressive subjective elements. In addition to this, the authors also annotate the source of the private state or event which is the speaker/writer in case of speech events and the experiencer in

case of the private state. (Wilson & Wiebe, 2005) extend this annotation scheme with the attributes – “Attitude” and “Target”, which respectively, denote the stance taken by the source and the entity about which the private state is expressed.

The work of (Read, Hope, & Carroll, 2007) is primarily focused on annotating opinions or appraisal units (attitude, engagement and graduation) in texts, which share similar notion with the Appraisal Framework developed by (Martin & White, 2005). Other works on annotating texts include (Carlson, Marcu, & Okurowski, 2001), (Taboada & Renkema, 2008) etc. which deal with text annotation in the discourse level employing discourse connectives and discourse relations.

However, despite these efforts, the development of a suitable annotation scheme for corpus annotation from the perspective of analyzing opinion and argumentation hand in hand in opinionated texts seem to be clearly missing. While the existing annotation schemes and guidelines may be sufficient for annotating private states, appraisal units, discourse units and even possibly some rhetorical relations, for analyzing the argumentation structure, it is necessary to annotate the thesis statements and the support statements, determine the orientation of the supports (either “For” or “Against”) with respect to the thesis statement, and identify the discourse relations between them. In addition to this, it is also necessary to determine the commitment or intent levels of the opinions and the overall persuasion effects in opinionated texts. This then requires us to make some additional provisions in the annotation scheme which are as follows:

- i) Introduction of some metadata of the source text like date and source of publication useful for source attribution in opinionated texts.
- ii) .Attributes for identifying arguments and their components as well as for determining the orientation of the supports.
- iii) Attributes for determining the strength of opinions and arguments or commitment level and persuasion effects in opinionated texts.

Considering the above issues and after a manual analysis of selected opinionated texts from the corpus as mentioned in the previous section, I develop a semantic tagset specifically designed

for the annotation of opinion and argumentation elements in opinionated texts; the tagset with brief explanation of the tags is provided in Table 3-3 below:

Table 3-3: Semantic tagset

Parameters	Possible values/Explanation
Topic	The title or topic of the opinionated text
Gist	The summary or abstract of the opinionated text. Usually, this is provided in the form of one or more sentences at the beginning of each text.
Author	The name of the author if available. Generally, in editorials, the name of the author is not provided but in the case of op-eds ⁴ , usually the names of the author(s) are mentioned.
url	The uniform resource locator or the web link to the opinionated text.
Date	The date of publication of the opinionated text.
Source	The source or the news portal from where the opinionated text is taken from.
argument_id	The argument's identity number. For simplicity, in this annotation scheme, each

⁴ <http://www.merriam-webster.com/dictionary/op-ed>

	<p>paragraph is regarded as an argument. This is because in argumentative text, the basic rule is that a paragraph generally sticks to a particular idea with several supporting/refuting evidence to the given idea. The numbering of the argument starts from 0 and this increases globally in the whole text as the paragraph advances from top to bottom.</p>
statement_id	<p>The statement/sentence number within an argument or paragraph. Each sentence is considered to be a statement. The numbering of the statement starts from 0. The numbering of the statement is relative to each paragraph.</p>
statement_type	<p>Can be either an “Opening-statement-thesis” or “Support” but not both. Usually, an Opening-statement-thesis puts forward a claim or a belief and the Support supports or refutes the claim.</p>
support_type	<p>A supporting statement or sentence can take either of the three values – “For” or “Against” or “Neutral”. If the supporting statement supports the claim, it is said to be providing a positive support or “For” and if the supporting statement refutes the claim, it is said to be providing a negative support or “Against”. Similarly, if the supporting statement does not</p>

support or refute the claim, it is said to be “Neutral” with respect to the claim.

exp_type

A statement or sentence as an expression can take either of the three values – “Opinion”, or “Fact” or “Undefined”. A statement is tagged as an Opinion if it represents a view, emotion, judgment etc. Similarly, a statement is tagged as Fact if it expresses some factual information. If a statement cannot be tagged as an “Opinion” or a “Fact”, it is tagged as “Undefined”. Often, there may be situations, whereby a portion of the statement represents a Fact while the other portion is an opinion. However, currently, we handle just statements with either factual or opinionated expressions but not both.

opinion_orientation

If a statement or sentence has been tagged as “Opinion”, the attribute “opinion_orientation” can take either of the four values – “Positive”, “Negative” or “Neutral” or “Undefined”. There can be one or multiple opinion terms of different polarity or orientation in a statement but the statement has to be tagged taking into consideration the overall effect in terms of opinion orientation. If the statement does not bear any particular opinion orientation, i.e., either “Positive” or “Negative”, it is tagged as

“Neutral”. Similarly, if a statement cannot be classified into any one of the three types, it is tagged as “Undefined”.

opinion_strength

This attribute tags a statement or sentence for the overall opinion strength across seven extended scale parameters – “Lowest” or “Lower” or “Low” or “Average” or “High” or “Higher” or “Highest”. The general base strength categories is however, “Low”, “Average” and “High” with the other four grades resulting when one or more intensifiers or pre-modifiers come in front of the three base strength categories. A statement can have multiple opinion terms of varying strengths but the overall opinion strength has to be considered.

persuasion_effect

This attribute tags a statement or sentence with one of the values- “Yes” or “No”. If the sentence or statement has an overall persuasion effect or is of convincing nature, the attribute “persuasion_effect” takes the value “Yes”, otherwise, it takes a “No” value. The persuasion effect is determined by the presence or absence of words and expressions that are persuasive in nature.

Conditional	This attribute tags a sentence or statement with one of the values “Yes” or “No”. If the statement is of conditional nature, the attribute “conditional” takes the values “Yes”, otherwise, it takes a “No” value.
commitment_level	This attribute tags a statement or sentence with one of the values – “Low” or “Average” or “High”. The major decision to tag the sentences with one of the above values is determined by the presence of different modal and/or reporting verbs of varying commitment or intent level.
rhetorical_relation	This attribute tags the Support statement or sentence with one of the following values – “Elaboration”, “Result”, “Reinforcement”, “Contrast”, “Cause-Effect”, “Exemplification”, “Conditional” and “Explanation”. The tagging for the given attribute is based on the explicit discourse markers or connectives present in the Support statement with respect to the Opening-statement-thesis or in between the preceding or following Support statements with respect to the current Support statement.
rhetorical_relation_to	This attribute tags the statement or sentence with the respective identity number

establishing the rhetorical relation between statements or sentences. Generally, with an argument or paragraph, this attribute establishes the relation between the current statement and/or the preceding or following statement.

3.4. Manual annotation of the texts

Using the semantic tagset described in the previous section, the same set of 100 opinionated texts from the corpus initially used for analyzing the structure was manually annotated. The texts range from editorials to op-eds.

The purpose of the manual annotation was two-fold, the first being to study the coverage and adequateness of the current semantic tagset for annotating opinions and argumentation elements in real opinionated texts and the second being to evaluate whether human annotators did the annotation task with substantial agreement or not. In case of substantial agreement between human annotators, it can be implied that the machine can also do the job considerably well and in the opposite case, the task is even difficult for humans implying that it is much harder to achieve the same with a machine.

For illustration purpose, a section of a manually annotated file is provided as a sample in Figure 3.1 below. More samples of the manually annotated texts are presented in Appendix C.

```

<Argumentation topic="Policies and Programs" gist="" author="" url="http://ekantipur.com/the-kathmandu-post/2008/09/11/editorial/policies--programs/160182.html" date="2008/09/11" source="The Kathmandu Post">
<Argument id="0">
<Statement id="0" type="Opening-statement-thesis" exp_type="Fact">President Ram Baran Yadav on Wednesday presented the Maoist-led government's policies and programs for the fiscal year 2008/09 at the parliamentary session of the Constituent Assembly.</Statement>
<Statement id="1" type="Support" type_support="Neutral" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Elaboration" rhetorical_relation_to="0">The 90-point document puts emphasis on taking the peace process to a logical end.</Statement>
<Statement id="2" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration" rhetorical_relation_to="1">For this, the government has pledged to complete the drafting of the new constitution within the stipulated time, integrate Maoist combatants into the security forces in six months and push forward the process of socio-economic transformation.</Statement>
<Statement id="3" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="Low" rhetorical_relation="Elaboration" rhetorical_relation_to="2">The document speaks about institutionalization of new federal republic, though there are differences among the ruling parties on the modality of federal structure.</Statement>
<Statement id="4" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="No" conditional="No" commitment_level="Average" rhetorical_relation="Elaboration" rhetorical_relation_to="3">Among others, the most important part of the document is that the Maoists have invited the armed-groups active in the tarai for a dialogue.</Statement>
<Statement id="5" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="No" conditional="Yes" commitment_level="High" rhetorical_relation="Explanation" rhetorical_relation_to="4">Maoists seem to have realized that dialogue is the only way to resolve all the differences.
</Statement>
</Argument>

```

Figure 3-1: Sample of the manual annotation

It was found that the tagset was adequate in terms of tagging the opinionated texts for opinions and argumentation elements as the annotators did not have any complaints on the tagset for the annotation task. However, a major difficulty faced in the task was while tagging statements consisting of opinions of varying orientations and strengths as one had to decide on the overall opinion orientation and strength to the statement based on the contents of the statement. Below, in Table 3-4, I demonstrate a few representative examples from the corpus and the strategies adopted to come to a conclusion in terms of determining the overall orientation of the statements. The tags “pos” and “neg” respectively refer to positive and negative opinion orientations of the individual word(s) and phrase(s). Every opening tag <> has a closing tag </> and scopes of each opening tag is defined by the nearest closing tag. The use of the “pos” and “neg” tags is just for illustration purpose in terms of showing how the overall opinion orientation of the statement is computed. In the actual annotation, this is not practiced; rather the overall orientation is determined by reading the statement of the text directly. The overall opinion orientation of the statement is denoted by “Positive” or “Negative”.

Table 3-4: Illustration of the annotation process

S.No.	Statement	Remarks
-------	-----------	---------

1	<p>Bhattarai's <neg>biggest challenge will be <neg>to meet <neg>conflicting demands</neg> for resources</neg> <u>while at the same time</u> <pos>"leapfrogging" development</pos>, <pos>keeping the <pos>sound macro-economic parameters</pos> that he inherited on track</pos>, and <pos>giving 29 million people <pos>new hope</pos> for a <pos>better tomorrow</pos></pos></neg>. Overall opinion orientation: <Negative></p>	<p>The construct “while at the same time” gives continuity to the discourse of “biggest challenge”, which is negative.</p>
2	<p><neg><pos>The <pos>success</pos> of the <pos>peace process</pos></pos> <neg>will remain <neg>incomplete</neg></neg></neg><u>until</u> <pos><pos>some kind of <pos>solution</pos></pos> with the <neg>armed groups</neg> is found</pos> – be it through <pos>dialogue</pos> or <pos>prompt action</pos>. Overall opinion orientation: <Negative></p>	<p>The construct “until” puts a condition and gives continuity to the previous discourse of the peace process being incomplete, which is negative.</p>
3	<p>We urge that the diplomatic missions of the two countries <pos>to launch a <pos>positive initiative</pos></pos> <u>instead of</u> just <neg>refuting the news</neg> that comes in the media in order <neg>to hide their own <neg>incompetence</neg></neg>. Overall opinion orientation: <Positive></p>	<p>The construct “instead of” contrasts the opinion orientation of the latter clause, thus giving an overall positive orientation of the sentence.</p>

4	Not only that, the citizenry has <pos>acted smartly</pos> <u>and</u> <neg>sent the <neg>old and tainted forces</neg> packing</neg>. Overall opinion orientation: <Positive>	An embedded negative expression within a negative expression results in a positive orientation. See “sent the old and tainted forces packing”.
5	The leaders must see <pos>elections</pos> as<pos><pos>a chance to <pos>mend ways</pos></pos> <u>and</u> <pos>improve their <neg>tarnished reputation</neg></pos></pos> among the public.	The opinion orientation tag with the outermost scope gives the final opinion orientation. See “improve their tarnished reputation”.

For the annotation task, three researchers from the Opinion Mining research group with a fairly good knowledge and understanding of the English language were involved, one was a PhD student and the other two were MS students. Prior to the annotation work, they were provided some pre-training and orientation on the annotation work. In Table 3-5, an illustration of the inter-annotator agreement is presented for a sample annotation. To simplify things, only the “opinion_orientation” attribute for annotation has been taken as an example.

Table 3-5: Inter-annotator agreement for the attribute “opinion_orientation” on a text sample

Statement	Annotator 1	Annotator 2	Annotator 3
With the beginning of the New Year, the country is going to witness a new political environment.	Positive	Positive	Positive
The successful holding of the Constituent Assembly (CA) polls has put every Nepali in a very optimistic mood.	Positive	Positive	Positive
But the political challenges ahead are certainly daunting.	Negative	Negative	Negative
The CA to be convened after the election results are released will have to seal the fate of the monarchy once and for all.	Negative	Negative	Negative
The 240-year-old institution will be consigned to the political history of Nepal.	Negative	Negative	Negative
The country, which will soon be declared a republic, should not be left in a lurch.	Positive	Positive	Positive
The new representatives will have to execute the promises made at different times when several agreements were signed, beginning with the 12-point pact.	Positive	Neutral	Neutral

The most difficult task is federalizing the country.	Negative	Negative	Negative
The nature of federal Nepal, demarcation of the internal boundaries and sharing of power between the federal units and the center are some of the greatest challenges.	Negative	Negative	Negative

In Table 3-6 and Table 3-7, the frequencies of agreements and the computed figures for observed agreement A_0 (the number of instances in which annotator pairs agree divided by the total number of instances) and agreement by chance A_e . For illustration purpose, only the inter-annotator agreements between Annotator 1 and Annotator 2 have been shown.

Table 3-6: Pair-wise inter-annotator agreements in the form of a contingency table

	Annotaor1	Annotator 2				Total
		Positive	Negative	Neutral	Undefined	
Positive	3	0	1	0		4
Negative	0	5	0	0		5
Neutral	0	0	0	0		0
Undefined	0	0	0	0		0
Total	3	5	1	0		9

Observed agreement $A_0 = 3/9 + 5/9 = 0.333 + 0.555 = 0.888 = 89\%$

Table 3-7: Observed agreements and agreements by chance (contingency table with proportions)

Annotator1	Annotator 2				Total
	Positive	Negative	Neutral	Undefined	
Positive	0.33	0	0.11	0	0.44
Negative	0	0.60	0	0	0.60
Neutral	0	0	0	0	0
Undefined	0	0	0	0	0
Total	0.33	0.60	0.11	0	1

Agreement due to chance $A_e = 0.33 \times 0.44 + 0.60 \times 0.60 = 0.50 = 50\%$

Inter-annotator agreement measure, Cohen's kappa (Cohen, 1960):

$$K = (A_0 - A_e) / (1 - A_e) = (0.88 - 0.50) / (1 - 0.50) = 0.38 / 0.50 = 76\%$$

The inter-annotator agreement scores or kappa measure for other attributes were also more or less similar. The pair-wise inter-annotator agreement scores or kappa measure for some major tagging attributes are presented in Table 3-8 below:

Table 3-8: Inter-annotator agreement scores for some major tagging attributes

Tagging attribute	Inter-annotator agreement scores or kappa measure (K)		
	Annotator 1-2	Annotator 2-3	Annotator 1-3
exp_type	80%	82%	78%

opinion_orientation	76%	100%	76%
opinion_strength	77%	79%	77%
commitment_level	85%	86%	87%
rhetorical_relation_type	68%	70%	69%

The results in Table 3-8 indicate that there was a good amount of agreement between annotators in terms of using the given tagset for the annotation purpose. In general, the disagreements were noted for the attributes – “exp_type”, “opinion_orientation”, “commitment_level” and “rhetorical_relation_type”. The disagreements were resolved by mutual discussions as well as consultation with experts.

Other works on annotation on similar domains also have alike inter-annotator agreements. For example, in the work of (Wilson & Wiebe, 2003), they report the agreement as being 0.87. In the work of (Hope & Carroll, 2007), the authors report the agreements as a mean F-Score of 0.698 at the most abstract level through to 0.395 at the most concrete level. This tells us that the agreements greatly vary between the tasks and the nature of the texts being annotated.

Since the human annotator agreement defines the upper limit on the ability to measure automated performance of NLP systems (Resnik & Lin, 2010, p. 276, section 2.5), a considerable inter-annotator agreement implies that the performance of the machine in terms of automatic annotation can also be effectively measured as the human annotations can be used as a bench mark or gold standard data, a term popularly used in NLP to refer to data developed by humans, for comparison of automatically annotated data.

3.5. Conclusion

In this Chapter, the discussion started with the motivations behind building a corpus of opinionated texts for opinion and argumentation analysis. Then I gave an account of the corpus that was developed for this purpose. Further, the findings on the argumentation structure of opinionated texts based on a manual analysis of a few representative samples of texts from the corpus were presented. Then, we looked into the challenges from the perspective of opinion and argumentation analysis. To find partial solutions, a semantic tagset was then proposed as a basis for annotating texts as a resource for analyzing and synthesizing opinions and arguments in opinionated texts. Finally, a study of the inter-annotator agreement and the results of the study were presented.

What we have seen from the manual annotation process and the description of the semantic tagset is that an extensive amount of linguistic knowledge is required for an accurate annotation of these resources even for humans, let alone the machine. From this perspective, the way forward is going for the development of linguistic resources (lexicon, opinion patterns and contextual rules) that would aid the machine to accomplish the same task with reasonable accuracies but in a considerable short time. In the next Chapter, I will be talking about the development of linguistic resources, the challenges, the methodology of development and their essence from the point of view of opinion and argumentation analysis in opinionated texts.

CHAPTER 4 LINGUISTIC RESOURCES FOR AUTOMATIC ANALYSIS

In Chapter 3, we manually analyzed the opinions and arguments from a number of texts from the corpus of opinionated texts and came up with a framework for analyzing the texts via annotations using the semantic tagset. However, we realized that the task requires an extensive knowledge of linguistic and world knowledge even for humans, let alone the machine.

In this Chapter, we will look into the methodology of incorporating this knowledge to the software by developing the resources and evaluating their use and effectiveness in the context of existing resources, methods and techniques, current results produced by their use and limitations.

4.1. Sentiment/Polarity Lexicon

Sentiment/Polarity Lexicon is a valuable resource for determining the orientation or polarity of opinions in opinionated texts, particularly in the lexical, phrase and sentence levels. A few of such lexicons already exist for the English language, for example, the opinion lexicon developed by (Hu & Liu, 2004), (Liu, Hu, & Cheng, 2005), subjectivity clues developed by (Riloff & Wiebe, 2003), SentiWordNet developed by (Esuli & Sebastiani, 2006), (Baccianella, Esuli, & Sebastiani, 2010).

The opinion lexicon by (Hu & Liu, 2004) and (Liu, Hu, & Cheng, 2005) consists of 2041 positive terms and 4818 negative terms. Similarly, the subjectivity clues list developed by (Riloff & Wiebe, 2003) consists of clues arranged such that one line typically represents one subjectivity clue. There are altogether 7629 clues and each clue has been annotated with six fields – “type”, which takes either strongsubj(strong subjective) or weaksubj(weak subjective), “len”, which gives the length of the clue word, “word1”, which marks the clue itself, which can either be a root or stem form of the word or inflected or derived form of the clue, “pos1”, which is used for providing the parts of speech value to the clue, “stemmed1”, which takes the values, ‘yes’ or ‘no’ depending upon whether the clue is a stem word or not and finally the field “priorpolarity”, which marks the prior polarity of the clue, i.e., the initial polarity of the clue when their contexts of appearance in texts are not taken into consideration. Similarly, the SentiWordNet by (Esuli & Sebastiani, 2006)

and (Baccianella, Esuli, & Sebastiani, 2010).is based on WordNet version 3.0 and provides positivity and negativity scores to the synset terms of WordNet. The SentiWordNet has the synsets annotated for six fields, respectively, the “POS” referring to the parts of speech, “ID” referring to the identity number, “PosScore” representing the positivity score, “NegScore” denoting the negativity score, “SynsetTerms” referring to the synset term with sense number and finally the “Gloss” which provides an explanation to the synset term with examples. The pair (POS and ID) uniquely identifies the WordNet synset. The objectivity score for a particular synset term can be calculated by using the formula, $\text{ObjScore} = 1 - (\text{PosScore} + \text{NegScore})$.

The above mentioned lexicons are certainly useful resources but they solely do not seem to be appropriate for the task of opinion and argument analysis in opinionated texts because first and foremost, the polarities of words and even the strength and other parameters like subjectivity and objectivity would necessarily change depending upon the contexts of occurrence. Besides, the lexicons above have been developed for certain restricted domains like restaurants and movie reviews and hence they are not exhaustive of the opinion terms as far as sentiment terms from other domains are considered. Out of the above mentioned lexicons, I took the opinion lexicon by (Hu & Liu, 2004) as the baseline resource. The lexicon was found to be in general quite useful and effectively helped in determining the opinion bearing words and their orientation or polarity but it was found that the resource quickly breaks down when terms from the socio-political domain are encountered. Even the frequent terms like “treaty”, “pact”, “truce”, “agitation”, “mutiny”, “salvage”, “concensus”, “epidemics”, “brotherhood”, “bandh” etc. in the socio-political domain seem to be missing in the opinion lexicon. This served as the primary motivation to develop a separate sentiment polarity lexicon comprising of prototypically positive and negative terms, specifically from the corpus. I started with a small collection of 29 positive terms and 73 negative terms from the corpus. These terms were collected by a manual analysis of some random texts from the corpus. Further, consulting the online and available electronic resources like dictionaries, thesauri and the WordNet, the list was extended by adding some synonyms, inflected and derivational form of the words. A sample of the Sentiment/Polarity Lexicon is presented in Table 4-1 below. A more detailed list is presented in Appendix E.

Table 4-1: A sample of the Sentiment/Polarity Lexicon

Positive	Negative
right: proper, correct, ok, okay	sack: fire, terminate, discharge, dismiss
reform: reforms, reformed,	insubordinate: insubordination
democracy: democratic, democratized	defy: disobey, defiance
contribute: contributed, contribution	unilateral: unilaterally
hope: hopeful, hoping	withdraw: withdrew, withdrawal
thank: grateful, gratitude, thankful	hate: hated, hatred
respect: honor, dignity. dignified, respectful	damage: damaging, damage
integrate: unite, unity, united, integrated, integration, merge	contradict: contradiction, contradicting
salve: salvage, save	insurgent: insurgency
glory: glorious, famous	refuse: refusal, denial

Such a collection allows us to have a rich lexicon of wider coverage comprising of both domain-specific terms from the corpus and the domain independent terms from online resources. Currently, the Sentiment/Polarity Lexicon contains about 300 positive terms and 800 negative terms. The task of opinion and argument analysis in opinionated texts involves assigning an opinion label – Positive or Negative or Neutral or Undefined to each statement/sentence. To illustrate the use of the Sentiment/Polarity for the assignment of the opinion label in the sentence

level, an excerpt of a real text from the corpus and its corresponding analysis is presented in Figure 4.1 below. For ease of illustration, the text has been segmented in the sentence level and analyzed for opinions in the lexical and phrase levels. While the opinion phrases have been annotated in XML like tagging notation, the opinion words/expressions have been underlined.

TITLE@Maoists' double standard

DATE@2007 May 05

#URL@<http://ekantipur.com/the-kathmandu-post/2007/05/05/editorial/maoists-double-standard/108572.html>

- 1) A report of the UN Office of the High Commissioner for Human Rights in Nepal (OHCHR-Nepal), issued last week, manifests the <neg>glaring facts</neg> about the CPN-Maoist. {Overall orientation: Negative}
- 2) In the report the OHCHR-Nepal has starkly said that the Maoist cadres <neg>aren't complying</neg> with their party's commitments and <neg>are not respecting</neg> the rights of the Internally Displaced Persons (IDPs) to voluntarily and safely return home. {Overall orientation: Negative}
- 3) The UN rights body has also expressed a <neg>serious concern</neg> over the <neg>continued existence of the Maoists' parallel government and their Kangaroo courts</neg>. {Overall orientation: Negative}
- 4) The scrutiny by the UN rights body conspicuously underlines a strong message that the Maoists <neg>haven't abandoned the culture of intimidation</neg>. {Overall orientation: Negative}
- 5) The former rebels' <neg>non-committal attitude towards their promises</neg> seems to be propelled by their <neg>instantaneous fear</neg> that the people, who <neg>suffered in the hands of the Maoists</neg>, will <neg>claim the price of their pain</neg> once <neg>the rebels lose their control</neg>. {Overall orientation: Negative}
- 6) <pos>Equally fair observation</pos> is that the Maoists want to <neg>continue political vacuum</neg>in the villages. {Overall orientation: Negative}
- 7) For them, <neg>the return of the IDPs</neg> means <neg>the return of other political parties

Figure 4-1: Excerpt of the analyzed text from the corpus

The analysis presented in Figure 4-1 provides us some interesting observations on account of opinionated texts. I note some of the observations below:

i) Negation and polarity shifters

The auxiliary verb forms of the verb “to be” accompanied by “not” in full or contracted forms shift the polarity of the word/phrase opinions from one direction to the other. For example, “aren’t complying”, “are not respecting”, “don’t have the right” as shown in sentences 2 and 8 respectively have negative opinion orientations because of the presence of the negative construct “not”. Other similar negation constructs are “never”, “hardly ever”, “seldom” etc. and non-assertive words like “any”, “anybody”, “anything” etc.⁵ Works of (Wiegand, Balahur, Roth, Klakow, & Montoyo, 2010), (Davdar, Hauff, & Jong de, 2011), (Heerschop, Iterson, Hogenboom, Frasinca, & Kaymak, 2011) deal with negation in Sentiment Analysis.

ii) An embedded negative expression within a negative clause results in a positive expression

For example, the clause “don’t have the right to obstruct political activities” has a positive opinion orientation. Here “obstruct” comes within the larger expression “don’t have the right to...”

iii) Opinion expression units in texts

It was observed that the analysis essentially occurs in the phrase level and higher for coming up with any meaningful interpretation of the opinions being expressed in the sentence level. This justifies the need for developing rules for capturing

⁵ <http://englishgrammar.org/negative-verb-forms/>

opinion expression patterns, which necessarily involve phrases. Opinion expression patterns are discussed in the next section.

4.2. Opinion Bearing Terms and Patterns from the Corpus

In addition to the terms in the Sentiment/Polarity Lexicon discussed in the earlier section, bigrams also serve as good candidates for opinion bearing terms.. The bigrams typically represent the opinion bearing phrases from opinionated texts. In Table 4-2, a sample of the extracted bigrams from the corpus is presented. A more detailed list of the extracted bigrams from the corpus is presented in Appendix G. Similarly, by manually analyzing the texts from the corpus, some opinion patterns also have been developed. These opinion patterns are expressions comprising one or more words and/or idioms and phrases that come either together with parts-of-speech categories or in isolation. Such generalized opinion patters help us to correctly detect opinions in texts and at the same time classify the polarity or orientation and strength of opinions. In Table 4-3, a sample of the positive and negative opinion patterns from the corpus is presented.

Table 4-2: Sample of the extracted bigrams from the corpus

Extracted bigrams		
al-Quaeda-run city	bird flu	corruption-plagued Nepal
anti-American sentiments	blue-collar jobs	cost-effective intervention
anti-Putin protests	call-bypass syndicates	coup d’etat
bankrupt regime	chaotic traffic	dead constitution
Bhutanese democracy	CIA-engineered coup	devastatingly successful

Table 4-3: Sample of the Positive and Negative Opinion Patterns from the Corpus

Positive Opinion Patterns and Terms	Negative Opinion Patterns and Terms	Example
	Continue to+ Verb (+ infinitive + Negative)	The Maoists continue to disrupt the proceedings of the Constituent Assembly.
	Continue + Noun (+ Negative)	The YCL continued its excesses and extortions.
	Turn a deaf ear	
	Turn a blind eye	
	So called + Noun	So called democrats So called rightists So called leaders
Verb + Adverb (+ Positive)		Behave responsibly Do well Perform well
Height of + Noun (+Positive)		Height of success
	Height of + Noun (+Negative)	Height of anarchy Height of impunity

Symbol of +Noun		Symbol of peace
(+ Positive)		Symbol of unity
Symbol of +		Symbol of oppression
Noun (+ Negative)		
Negation	+	Not object
Negative verb		If the Maoists do not run away from the elections
Positive Adjective		Progressive agenda
+ Noun		Moderate stance
Anti-+(Noun/Adjective)		Anti-Maoists
		Anti-national

4.3. Intensifier/Pre-modifier Lexicon

For the task of analyzing the opinions and arguments in opinionated texts, besides determining the subjectivity (whether a given expression is an opinion or not) and detection of the orientation or polarity of opinions, it is necessary to assess the strength or degree or intensity of opinions. Adjectives and adverbs have a significant role in the determination of the strength or degree of opinions as they necessarily change the intensity or degree of opinions being expressed (Hatzivassiloglou & McKeown, 1997), (Hatzivassiloglou & Wiebe, 2000), (Chesley, Vincent, Xu, & Srihari, 2006), (Benmara, Cesarano, Picariello, Reforgiato, & Subrahmanian, 2007). Although, there can be finer grades of opinions, I have limited the grading to seven broad scales – “Lowest”,

“Lower”, “Low”, “Average”, “High”, “Higher” and “Highest” for my task. This corresponds to a scale within the range -3 to 3, where the mapping of the degrees to numeric values are as follows:

Lowest =-3, Lower =-2, Low =-1, Average = 0, High=1, Higher =2, Highest =3

The mapping above is partly guided by the three degrees of adjectives in English, viz., positive, comparative and superlative. In my case, positive degree refers to “Low”, comparative degree to “Average” and superlative to “High”. These three scales have been considered as the base strength categories. The remaining four scales “Lower” and “Lowest” and “Higher” and “Highest”, respectively on the “Low” and “High” sides are produced as a result of the possible occurrences of intensifiers and pre-modifiers in front of the three major degrees of adjectives – “Low”, “Average” and “High”. A few examples of the three degrees of adjectives from the corpus are shown below:

high, low, good, bad, few, wealthy, powerful, successful: **positive degree (“Low”)**

higher, lower, better, worse, fewer, wealthier, more powerful, more successful :
comparative degree (“Average”)

highest, lowest, best, worst, fewest, wealthiest, most powerful, most successful :
superlative degree (“High”)

In addition to adjectives, intensifiers and pre-modifiers also have been considered for the determination of the different degrees of strength of opinions. Intensifiers are essentially adverbs which are reported to have three different functions – emphasis, amplification and downtoning. Pre-modifiers, on the other hand, come in front of adverbs and adjectives. Both intensifiers and pre-modifiers play a role in conveying a greater and/or lesser emphasis to do something. The intensifier lexicon is provided in Table 4-4 below. The contents of the lexicon (type and value) have been adopted from (Capital Community College, 2014).

Table 4-4: Intensifier lexicon

Type	Value	Occurrences from the corpus
Emphasizer	Really: truly, genuinely, actually	This is really a good idea.
	Simply: merely, just, only, plainly	I simply cannot say.
	Literally	I would literally trust his judgments over mine.
	For sure: surely, certainly, sure, for certain, sure enough, undoubtedly	All we can say for sure at this point is...
	Of course: naturally	There were many tactical and strategic compromises along the way, of course .
Amplifiers	Completely: all, altogether, entirely, totally, whole, wholly.	Men and women are completely equal in value and dignity.
	Absolutely: totally, definitely, without, question, perfectly, utterly.	I just told them that we should be absolutely quite.
	Heartily : cordially, warmly, with gusto and without reservation	Heartily approve of socialism.
Downtoners	Kind of : sort of, kinda, rather, to some extent, almost, all but	The opponents were kind of satisfied with the answers of the Prime Minister.
	Mildly: gently	The Prime Minister mildly protested the proposal.

The entries in Column 2 of Table 4-4 highlighted in bold text represents the super set to which other sub-entries belong to. Below, we discuss the role of each category of intensifiers in terms of modifying the strength of opinions in example texts from the corpus:

Example 1: “The loss of the Corby bi-election is a really significant watershed”.

In the example text above, the intensifier (emphasizer) “really” emphasizes the adjective “significant”, thus increasing its intensity or degree to one level further up. In this respect, since the adjective “significant” represents the positive or “Low” degree, the intensifier “really” modifies the intensity of the strength of the adjective to “Average”.

Example 2: “The electoral commission was **absolutely** right to announce a review of the debacle”.

In the example text above, the intensifier (amplifier) “absolutely” amplifies the adverb “right”, thus increasing its intensity or degree to the highest level. In this respect, the intensifier “absolutely” modifies the intensity of the strength of the adverb to “Highest”.

Example 3: “Admittedly, this sounds rather disconcerting”.

In Example 3 above, the intensifier (downtoner) “rather” downtones the adverb “disconcerting” to one level down, thus modifying the intensity of the strength of the adverb to “Lower”.

Next, in Table 4-5, the pre-modifier lexicon and their contribution to the overall strength of opinion expressions is presented. The contents of the lexicon have been adopted from (Capital Community College, 2014).

Table 4-5: Pre-modifier lexicon

Adverb/Adjective (Initial strength)	Pre-modifier	Modified strength (Lowest, Lower, Low, Average, High, Higher, Highest)
--	---------------------	---

Fast (Low)	Very	Very fast (High)
Careful (Low)	Lot more	Lot more careful (High)
Better (Average)	Much	Much better (High)
Serious (Low)		Much much better (Higher)
		Much more serious (Higher)
Good (Low)	Somewhat	Somewhat good (Average)
	Quite	Quite good (average)

4.4. Report and Modal Verbs Lexicon

For the task of determining the strength of opinions and arguments in opinionated texts, it is also necessary to analyze the intent or commitment level of the statement under consideration with respect to some thesis statement. One way of doing this is by looking at the choice of report or modal verbs used in the respective statements. The higher the degree of assertiveness a modal/reporting verb is, the stronger the commitment or intent level of the statement. In Table 4-6, the modal verb lexicon and its role in determining the commitment or intent level is presented. The contents of the lexicon have been adopted from (Capital Community College, 2014).

Table 4-6: Modal verbs lexicon

Type	Verb	Strength effects (Low or Average or High)
Ability/Possibility	Can	Average
Ability/Possibility	Could	Low
Permission	May	Average
Permission	Might	Low
Advice/Recommendation/Suggestion	Should	Average
Necessity/Obligation	Must, Have to	High

Similarly, in Table 4-7, the reporting verb lexicon is presented. The contents of the lexicon have been adopted from (The University of Adelaide, 2009). I have re-interpreted their three categories, “weaker position”, “neutral position” and “stronger position” as “Low”, “Average” and “High” to better fit into the context of commitment levels conveyed by reporting verbs.

Table 4-7: Reporting verbs lexicon

Type	Low	Average	High
Agreement	admits, concedes	accepts, acknowledges, agrees	applauds, congratulates, praises, supports

Argument and Persuasion	Apologizes	assures, interprets, reasons	encourages, justifies, warns	alerts, argues, boasts, contends, convinces, warns
Believing	guesses, hopes, imagines	believes, declares, expresses	claims, asserts, insists, holds	guarantees,
Disagreement and questioning	doubts, questions	challenges, disagrees, questions	debates, accuses, complains	attacks,
Presentation	Confuses	comments, reports, states	defines, announces, promises	
Suggestion	alleges, intimates, speculates	Advises, posits, suggests	advocates, recommends, urges	

To illustrate the use of the intensifier and the pre-modifier lexicon as well as the report and modal verb lexicon for determining the intensity or degree of opinions and the commitment or intent level of the statements, an excerpt of the real text from the corpus and its corresponding analysis is presented in

Figure 4-2 below.

TITLE@Egypt and the Palestinian question

AUTHOR@Abdullah Al-Arian

DATE@05 Feb 2011

URL@<http://www.aljazeera.com/indepth/opinion/2011/02/20112512042978473.html>

- 1) Along with the laundry list of domestic grievances <commitment_level="Average">expressed</commitment_level>by Egyptian protesters <commitment_level="High">calling</commitment_level> for an end to the regime of Hosni Mubarak, the popular perception of Egypt's foreign policy has also been a focal point of the demonstrations.{Overall commitment level: "High"}
- 2) Signs and chants have <commitment_level="High">called</commitment_level>on Mubarak to <commitment_level="High">seek</commitment_level>refuge in Tel Aviv, while his

Figure 4-2 : Excerpt of the analyzed text for the strength of opinions and commitment level

4.5. Discourse Marker & Rhetorical Relations Lexicon

For analyzing the opinions and arguments in the sentence and higher levels, we need to determine the rhetorical or discourse or coherence relations. These relations are crucial in establishing relationships between passages of texts. As discussed in previous Chapters, discourse markers can serve as effective sign-posts to signal the presence of discourse or coherence or rhetorical relations in any discourse (Marcu, 1998; Fraser, 1999). In Table 4-8, the discourse markers and rhetorical relations lexicon is presented. The list of discourse markers have been adopted from (London Metropolitan University, 2010).

Table 4-8: Discourse markers and rhetorical relations lexicon

Rhetorical relations	Discourse markers
Elaboration	after, before, first, all the while, in the past,...
Result	briefly, hence, overall, thus, in brief, to end,...
Reinforcement	again, also, too, in addition, above all, most of all, ...
Contrast	against, instead, rather, still, versus, yet, even, so, ...
Cause-Effect	hence, since, therefore, thus, whenever, as a result, ...
Exemplification	indeed, namely, for example, in effect, such as, ...
Conditional	else, if, otherwise, unless, until, while, as long as, ...

Explanation	since, because, due to,...
--------------------	----------------------------

To illustrate the use of the Discourse Markers and Rhetorical Relations Lexicon in analyzing the discourse or coherence or rhetorical relations between supporting statements in texts, an excerpt of real text from the corpus and its corresponding analysis is presented in Figure 4-3, below.

TITLE@In praise of ... Jimmy Carter

DATE@2008 Apr 18

URL@http://www.theguardian.com/commentisfree/2008/apr/18/usa

<Rhetorical_relation="Exemplification">Like the Kennedy Library in Boston, where Gordon Brown makes the main foreign policy speech of his US visit today, most American presidential libraries are monuments to the past.</Rhetorical_relation> <Rhetorical_relation="Contrast">The Carter Centre, near Atlanta, is totally different.</Rhetorical_relation> <Rhetorical_relation="Exemplification">Like its begetter, Jimmy Carter, it is focused on the future.</Rhetorical_relation> <Rhetorical_relation="Reinforcement">The centre thrums with constant activity.</Rhetorical_relation> <Rhetorical_relation="Elaboration">Its slogan, "Waging peace, fighting disease, building hope" sums up the work of the most active ex-president the US has ever seen.</Rhetorical_relation>

Figure 4-3: Excerpt of the analyzed text for Discourse Marker and Rhetorical Relations

4.6. Automatic Annotation Tool

An automatic annotation tool was developed for automatically annotating the opinionated texts from the corpus for opinions and argumentation elements. The linguistic resources discussed so far in the earlier sections in this Chapter were incorporated into the tool. A high-level system architecture of the annotation tool is presented in Figure 4-4 below.

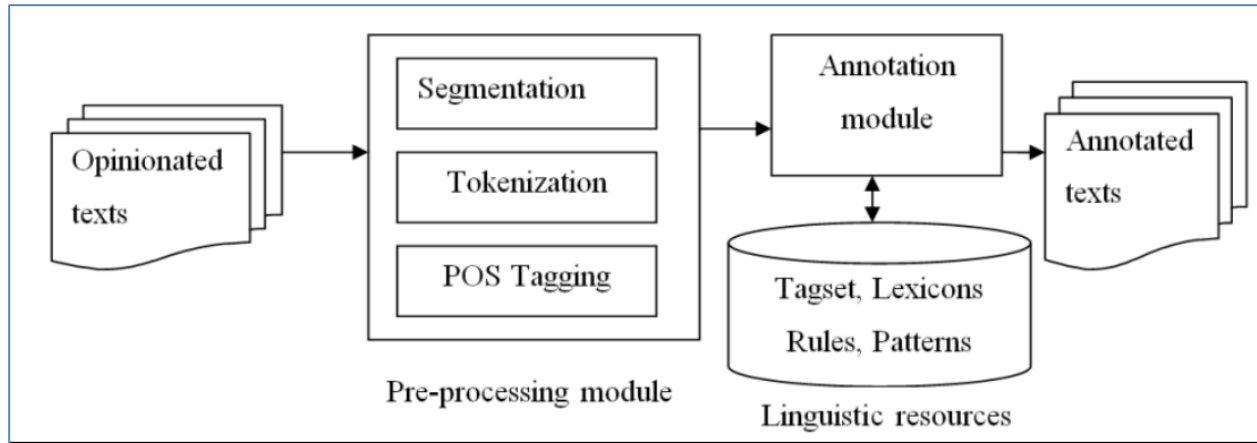


Figure 4-4: High-level system architecture of the annotation tool

As shown in Figure 4-4, the tool comprises two modules, respectively, the pre-processing module, the annotation module, and the linguistic resources database. Opinionated texts are passed to the pre-processing module, which performs segmentation of the texts into paragraphs and sentences. Then the texts undergo tokenization and POS tagging. The pre-processed texts are then passed to the annotation module, which annotates the opinions and argumentation elements in the texts by consulting the linguistic resources database, which comprises the semantic tagset, lexicons, contextual rules for tagging and the different opinion patterns for determining the polarity or orientation. The output of the tool is annotated texts for opinions and argumentation elements.

The performance of the automatic annotation tool was tested by annotating 100 randomly selected texts from the corpus. The expected output is tagging of the text in the sentence level, respectively for segmentation of the text into paragraphs and sentences, determination of the orientation and strength of opinions, commitment and intent level of the statements, rhetorical relation between statements etc. The accuracy of the performance of the tool was evaluated manually in terms of annotations by the machine compared to what a human would have annotated for the same. Since the annotation tool highly relies on the linguistic resources, we also carry out a comparative analysis of the use of the baseline linguistic resource (opinion lexicon by Hu & Liu, 2004) versus the extended linguistic resources(sentiment/polarity lexicon by (Hu & Liu, 2004) augmented with domain specific terms and patterns as well as contextual rules) for the same 100

texts mentioned above. The accuracy of the performance of the automatic tagger application in terms of tagging is calculated as follows:

$$\text{Accuracy} = \frac{\text{tag}}{T} \quad (4.1)$$

where T = Total number of tagged sentences
Total number of correctly tagged sentences and tag =

Currently, the annotation tool does not adequately handle the annotation of the attributes “Fact” and the support type (“For” or “Against” or “Neutral”). Hence, these tasks are not evaluated at the moment.

The accuracy scores for the different tagging tasks are presented in Table 4-9 below:

Table 4-9: Accuracy scores of the automatic annotation tool

S.No.	Annotation task	Accuracy (%)
1	Opinion orientation	61.5%
2	Opinion strength	63.75%
3	Commitment or intent level	72.5%
4	Rhetorical relations	47.5%

Similarly, in Table 4-10, the accuracies of the annotation task for the attribute “opinion_orientation” using the baseline resource and the extended linguistic resource are presented.

Table 4-10: Accuracy scores of the tool with the baseline and extended linguistic resources

S.No.	Annotation task (Opinion orientation) vs. linguistic resource	Accuracy
1	Baseline linguistic resource	55%
2	Extended linguistic resource	70%

The accuracy scores in Table 4-9 show that the tagging tasks (determining the orientation or polarity, strength of opinions and commitment or intent level) have been accomplished with fairly good results. The scores for each of these individual tasks are expected to further improve as the linguistic resources get further enhanced in terms of size and coverage. The task currently performing the least is determining the rhetorical relations. This is partly because implicit discourse markers are currently not considered as a potential sign-post for denoting the presence of rhetorical relations in between statements. The other possible reason could be the ambiguities posed by discourse markers with respect to the rhetorical relations that they signal, which has not been dealt with at the moment. It is expected that the performance of this particular task will also improve as some special tailored rules are designed to address the above mentioned situations.

Similarly, the accuracy scores in Table 4-10 show that the performance of the extended linguistic resource is much better than that provided by the baseline linguistic resource. This is understandable as the extended linguistic resource has a rich collection of domain specific terms and patterns from the opinionated text domain in addition to the opinionated terms from the baseline linguistic resource. It is expected that the accuracy scores for the opinion orientation tagging task will further improve as more of such domain specific terms and patterns are gathered.

4.7. Conclusion

In this Chapter, we looked into the different linguistic resources – Sentiment/Polarity Lexicon, Intensifier/Pre-modifier Lexicon, Report/Modal Verb Lexicon, Discourse Markers and Rhetorical Relation Lexicon, opinion bearing words, phrases and patterns from the corpus etc.

and illustrated how they can be used for annotating opinions and argumentation elements in real texts from the opinionated text domain. Then the high-level system architecture of the annotated tool was presented where we saw that the linguistic resources form a vital component of the system. Finally, we discussed the evaluation of the results of the tool in terms of the use of the linguistic resources for analyzing opinions and argumentation elements in opinionated texts. The results have shown that the annotation tasks and correspondingly the analysis of opinions and argumentation elements have been achieved with fairly good accuracies. The accuracy scores for the individual tasks are expected to improve as the linguistic resources are further enhanced in terms of coverage and size.

So far, we have seen how linguistic resources can be devised to analyze the opinion and argumentation elements in opinionated texts. However, we are yet to look into how these analyzed opinions and argumentation elements from different opinionated texts sources can be collected together in the form of a synthesis. This requires retrieving related texts\paragraphs from a single or multiple sources with respect to a common thesis statement or issue of discussion and then grouping the retrieved texts by some useful attributes like date of publication, source, opinion orientation etc. The next Chapter will deal with the tasks of opinion search, retrieval and synthesis.

CHAPTER 5 SYNTHESIS OF OPINIONS AND ARGUMENTS

Opinionated texts like op-eds, blogs, letters to editors etc. from online news sources represent rich sources of publicly available opinions and sentiments on events and happenings. Since the different news agencies do not necessarily share their views on a particular topic, the opinion content is usually disparate and guided by their pre-conceived socio-political ideologies and vested interests. Getting a consolidated view and perspectives of different editorials on the event is thus an interesting, yet a challenging research problem. The data collected via such research can help obtain a relatively true picture of events or happenings and potentially expose the inherent biases and prejudices. Moreover, such data facilitates time-event and thus trend analysis of the opinions expressed and stances taken by the different sources.

The task of synthesizing opinions and arguments with respect to the thesis statement in opinionated texts necessarily involves three sub tasks – i) retrieving the related texts from the corpus, ii) for the retrieved texts, extracting the opinion and argument information like fact or opinion, polarity and strength in case of opinions, support type, rhetorical relation, date of publication of the article, source or name of the publishing agency etc. from corresponding annotated texts and iii) finally construct a synthesis of the information again in the annotated format and/or some format easy for the users to understand. The work is related to (Zhang, Yu, & Meng, 2007) but in my work, the subjectivity and opinion polarity analysis of the retrieved texts is not conducted. Rather, the information of the retrieved texts (file name, line and paragraph number) is utilized to extract the relevant opinion and argumentation analysis from the matching annotated texts which correspond to the retrieved text fragments.

The given work is different from information retrieval (Salton & Lesk, 1968) in that here the task necessarily deals with texts/paragraphs within documents rather than the documents as a whole. Since in the given work, related texts/paragraphs are sought with respect to the thesis statement, it is necessarily the case that even texts/paragraphs within the same document can have varying degrees of text relatedness. That is, for each text/paragraph considered to be related with the thesis statement, an overall text relatedness score needs to be computed. This helps to filter the

highly related text fragments from the less related ones, thus reducing the unnecessary overhead of dealing with relatively less related texts while extracting the opinions and arguments from the annotated texts in the extraction phase.

In Chapter 4, the discussion was focused on the resources and techniques for analyzing opinions and argumentation expressed over a certain topic or event in a single opinionated text. In this Chapter, the discussion will be extended to synthesizing opinions and arguments from multiple opinionated texts. First of all, this Chapter will look into retrieving related texts/paragraphs in opinionated texts with respect to the thesis statement or issue of discussion. Following the discussion on retrieval of related texts, this Chapter will then look into how the information from related texts/paragraphs from multiple opinionated texts can then be used in conjunction with the annotations from the annotated texts for carrying out a synthesis of opinions and arguments.

5.1. Retrieving Related Texts from Opinionated Texts

For retrieving the related texts from the corpus, a query is formed based on the thesis statement, which is then passed to the text retrieval tool. Based on the query, the tool returns a set of related texts with their respective overall text relatedness scores. Below, I define some concepts associated with the problem of retrieving related texts/paragraphs from opinionated texts:

- i) A query expression (q) is essentially a list of words passed individually or in combination with different logical operators like AND, OR, NOT to the text retrieval tool as input. The query is based on the thesis statement (TS) or the issue of discussion but need not necessarily consist of exactly the same terms as in the latter.
- ii) A query term (q_t) of the query expression (q) may be an exactly matching term (q_{te}) to that of the thesis statement. For example, there may be a term “violent” in both the thesis statement and the text paragraph under consideration.

- iii) A query term of the query expression may be a related term (q_{tr}) with respect to the terms of the thesis statement. For example, there may be the term “chaos” in the thesis statement and “confusion” in the text paragraph under consideration.
- iv) A query term of the query expression may be a domain specific term (q_{td}) associated with the term of the thesis statement. For example, if the thesis statement were something like “2007 was a violent year full of post-conflict chaos and confusion” and the context of the statement is for Nepal, some probable domain-specific terms relevant to “Nepal” and “2007” would be “CA elections”, “Maoists”, “Constituent Assembly”, “YCL”, “Peace Process” etc. The derivation of the domain specific terms is explained in Section 5.2.
- v) The relatedness score (RS) of each term, for any of the three categories – exactly matching, related or domain specific is the weight of the term’s relatedness in the text/paragraph with respect to the thesis statement.
- vi) The overall relatedness score (ORS) of the text/paragraph is the aggregate weight of text/paragraph under consideration produced by adding the weights of the individual terms in the text/paragraph.

Now that we have defined the text relatedness concepts, let us try to look at the problem through an example.

2007 was a violent year full of post-conflict chaos and confusion.

Figure 5-1: Thesis statement – A year of hope, KTMPOST, Jan 02, 2008

The **year** that has just bade us farewell was one of **post-conflict chaos** and **confusion**. Both the two dates announced for the *constituent assembly elections* came and went without the *vote* taking place. The past **year** will be remembered for **violent movements**, especially in the *tarai*, and a few but serious conflicts in the western *tarai*. Crime and political murders spread fast and hard in the *tarai*. The *Young Communist League (YCL)* formed last **year** also set new records in violence and extortion. The **year** also saw the height of anarchy, impunity and lawlessness. The streets of *Kathmandu* gained infamy for the most anarchic traffic in the whole world. This was the **year** *Nepal* declared itself a *Federal Democratic Republic*, but the decision needs to be endorsed which will probably happen in 2008. In a nutshell, the year **2007** was a year filled with turmoil.

Figure 5-2: Text/Paragraph Sample 1 – A year of hope, KTMPOST, Jan 02, 2008

Nepal has been declared a *republic*, but it will only take formal effect sometime in the middle of next **year** after it is ratified by the *constituent assembly*. But the *king* is in his palace, still paid a salary by taxpayers' money.

Figure 5-3: Text/Paragraph Sample 2 – Peace train, Nepali Times, Dec 28, 2007 – Jan 03, 2008

Three different notations have been adopted to highlight the terms in the text – the exactly matching terms are highlighted in bold, domain-specific terms are italicized and related terms are underlined. Based on the matching terms, the relatedness of the individual texts/paragraphs with respect to the thesis statement is computed as follows:

$$RS(q_t) = \frac{f(q_t, P)}{T} \left(\ln \left(\frac{N}{df} \right) \right) \dots \dots \dots (5.1)$$

where, $RS(q_t)$ denotes the relatedness score of q_t , $f(q_t, P)$, the frequency of q_t in Paragraph P and T , the total word count in the document containing q_t . While computing the total word count, I do not consider stop words. A list of stop words that has been used for this purpose is presented in Appendix H. Similarly, N refers to the total number of documents in the collection or corpus and df denotes the number of documents where the term q_t occurs in the corpus. In equation 5.1, standard concepts, term frequency and inverse document frequency from information retrieval have been adopted, where $\frac{f(q_t, P)}{T}$ denotes term frequency and $\ln\left(\frac{N}{df}\right)$ denotes inverse document frequency.

The equation 5.2 below calculates the overall relatedness score:

$$ORS(TS, P) = \sum_{q_t \in \{q_{te}, q_{tr}, q_{td}\}} RS(q_t) \dots \dots \dots (5.-1)$$

where, $ORS(TS, P)$ denotes the overall relatedness score of TS and P and q_t can be either an exactly matching term q_{te} or a related term q_{tr} or a domain-specific term q_{td} . Similarly, N refers to the total number of documents in the collection or corpus and df denotes the number of documents where the term q_t occurs in the corpus.

5.2. Deriving Related and Domain-Specific Terms from the Corpus

For deriving the query terms (related and domain-specific) from the corpus, the word co-occurrences and the frequencies in the corpus have been exploited. The assumption here is that two words are related if they co-occur in the same text/paragraph, multiple times in different opinionated texts across the corpus. For this purpose, the relative or proportionate frequency of terms is calculated given the frequency of the occurrence of the terms from the thesis statement in the returned matching text or document. The relative or proportionate frequency of the terms is calculated by using the following formula:

$$RelFreq(w_2 | w_1) = \frac{df(w_1, w_2)}{df(w_1)} \dots \dots \dots (5.3)$$

where, $df(w_1, w_2)$ represents the number of texts or documents where the two terms co-occur whereas $df(w_1)$ represents the number of documents in the corpus where the term occurs.

In Table 5-1, the relative co-occurrence frequencies of the terms are presented. The observed frequencies of each term are shown within brackets, whereas the relative co-occurrence frequencies of the terms are placed within the cells in the table. While computing the frequencies of the words with slight grammatical/spelling variants, for example, the singular and plural cases (“Conflicts” and “Conflict”, “Elections” and “election”), I take the sum of the frequency counts together and not separately.

Table 5-1: Relative co-occurrence frequencies of the terms in the corpus

Thesis statement terms →	Violent(732)	Conflicts(283)	Confusion (206)
Related terms ↓			
Chaos (307)	=14/732=0.019=2%	=5/283=0.017=1.7%	=12/206=0.058=5.8%
Movements (375)	=51/732=0.069=7 %	=20/283=0.070=7%	=4/206=0.019=2%
Crime (889)	=56/732=0.076=7.6 %	=2/283=0.007=0.7%	=7/206=0.033=3.3%
Murders (92)	=6/732=0.008=0.8%	=1/283=0.003=0.3%	=1/206=0.004=0.4%
Extortion (68)	=6/732=0.008=0.8%	=1/283=0.003=0.3%	=1/206=0.004=0.4%
Anarchy (83)	=7/732=0.009=0.9%	=1/283=0.003=0.3%	=1/206=0.004=0.4%
Impunity (284)	=14/732=0.019=2%	=5/283=0.017=1.7%	=1/206=0.004=0.4%
Lawlessness (47)	=3/732=0.004=0.4%	=1/283=0.003=0.3%	=3/206=0.014=1.4%
Infamy (9)	=1/732=0.001=0.1%	=1/283=0.003=0.3%	=1/206=0.004=0.4%
Excesses (85)	=2/732=0.002=0.2%	=2/283=0.003=0.3%	=0/206=0%

In Figure 5-4, the figures from Table 5-1 are presented in the form of a bar chart diagram.

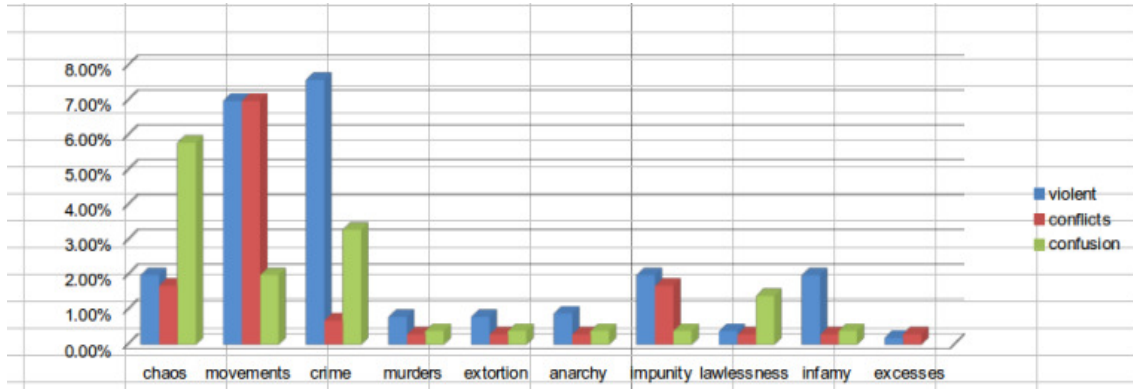


Figure 5-4: Relative frequencies of co-occurrences of terms in the corpus

We can see from the chart in Figure 5-4 that the words “violent” and “confusion” form their own sets of related terms, which can be organized as {violent: crime, movement, impunity, infamy, chaos, murders, extortion, anarchy, lawlessness and excesses} and {confusion: chaos, crime, movement, lawlessness, murders, extortion, anarchy, impunity, infamy, excesses}. The relatedness information can be used while forming related query terms to the exact terms from the Thesis statement.

Similarly, for deriving the domain-specific terms, the corpus is crawled for the top N most frequently occurring terms. Obviously, if contexts are involved, for example, “Nepal” and the year “2007-2008” in case of the example in Figure 5-4, we also need to keep this information into consideration. This means, we necessarily need to look into the opinionated texts from the year 2007 and 2008. In doing so, the stop words, which tend to occur frequently in texts are not included in the list. In addition to this, the frequency counts of words with close spelling/grammatical variants like “Nepali” and “Nepalis” or “Nepalese”, “Maoist” and “Maoists” etc. should be counted together as part of a single entry. The top 20 most frequently occurring terms in the corpus with their frequencies are provided in Table 5-2. A larger list comprising the top 200 most frequent terms from the corpus are provided in the Appendix F.

Table 5-2: Top 20 Domain-Specific Terms from the Corpus for the year 2007-2008

Term	Frequency				
2007	1052	People	800	Rights	480
2008	1038	Kathmandu	786	State	407
Maoists	998	Country	740	Leaders	356
Government	932	Political	651	Public	290
Party	867	Years	524	Democracy	256
Nepal	865	Elections	505	Process	251
		Nepali	499	CA	246

Similarly, a list of phrases and their frequency of occurrences in the corpus is provided in Table 5-3. The phrase list also serves as domain-specific terms from the corpus.

Table 5-3: Domain-Specific Terms (Phrases) from the Corpus for the year 2007-2008

Phrase	Frequency
Constituent assembly	193
Interim constitution	70
CA elections	67
Young Communist League	32
Constituent assembly elections	30
Seven parties	15
23-point agreement	7

Federal Democratic Republic	6
Mixed election system	3
Tarai movement	2
Madhesi militant factions	1

5.3. Experimental Results on Retrieval of texts

The GNU Grep tool⁶ was used as the text retrieval tool for retrieving the texts/paragraphs with respect to a particular thesis statement. The tool searches for lines containing matching patterns in files. The query, in this case, are words/phrases or regular expressions in the form of one or more “AND”, “OR”, “NOT” operators. Below in Table 5-4, a few example queries have been provided.

Table 5-4: Example queries

Query expressions	Explanation
Grep –inw “2007” Path Grep –inw “violent” Grep –inw “year”	Print the line matches in the files in the given path denoted by “Path” which match the words “2007” AND “violent” AND “year”. Option “i” ignores upper and lower case, “n” prints the matching line number of the text and “w” matches exact word.
eGrep –in “violent year” Path Grep –in “Nepal”	Print the line matches in the files in the given path denoted by “Path” which match the word “violent” or “year” and the word “Nepal”.

⁶ <http://www.gnu.org/software/grep/>

eGrep -in “violentlyear” Print the line matches in the files in the given path denoted by **Path | Grep -in “Nepal”** “Path” which match the word “violent” or “year” and the word “Nepal” found in the folder other than “2012”. Here the option “v” inverts the matching, i.e., prints the matching text in all folders except “2012”.

Let us now get back to our earlier example (See Figure 5-1 for the Thesis statement and Figure 5-2 and Figure 5-3 for the texts/paragraphs and examine the process of computing respectively the relatedness score of each matching term and the overall relatedness score of the texts/paragraphs. In the given example, $T = 282$ and $N=10807$, where they respectively denote the number of words in the text minus the stop words and the total number of documents in the collection or corpus. A working example is presented in Table 5-5 below:

Table 5-5: Computing the relatedness scores of query terms and overall relatedness scores for text/paragraph

S.No.	Type of query term	Frequency of query term	Relatedness score (<i>RS</i>)
1	Exactly matching	year(7)	0.015
2		conflict(1)	0.019
3		chaos(1)	0.012
4		confusion (1)	0.014
5		violent(1)	0.009
6		2007(1)	0.003
7	Related	conflicts(1)	0.008

8		crime (1)	0.008
9		political murders (1)	0.03
10		violence (1)	0.006
11		extortion(1)	0.017
12		anarchic (2)	0.027
13		impunity(1)	0.012
14		lawlessness(1)	0.019
15		infamy (1)	0.025
16		turmoil(1)	0.014
17	Domain-specific term	elections(1)	0.005
18		movements(1)	0.011
19		constituent assembly (1)	0.011
19		tarai (2)	0.026
20		Young Communist League	0.019
21		Kathmandu (1)	0.005
22		Nepal (1)	0.005
23		Federal Democratic Republic (1)	0.025

Overall score(<i>ORS</i>)	relatedness	0.345
--------------------------------------	--------------------	--------------

Next, in Table 5-6, the texts/paragraphs returned by the tool with respect to the thesis statement in Figure 5-1 along with their respective overall relatedness scores are presented in decreasing order.

Table 5-6: Texts/Paragraphs with their overall relatedness scores in decreasing order

Source	Text	Overall relatedness scores
<p>Paragraph 3:</p> <p>ktmpost/2008/01/02_year-of-hope_132912.txt</p>	<p>However, the fate of 2008 and beyond depends much on the behavior of our democratic political parties, the Maoists, army, civil society, media, et al. If we all behave responsibly, we will be able to hold the CA elections, accomplish the peace process, restart our economic motor and finally build a new Nepal. But it is definitely not an easy job. The major issue is, of course, the CA elections; but the problems of the tarai people, janajatis, women, oppressed, marginalized and the like are not so easy to address. More so when we are immersed hopelessly in the politics of corruption, nepotism, favoritism, cronyism and other sorts of ills. If the Maoists do not run away from the elections, if the recently formed and old parties in the tarai live up to their promise to allow the CA elections to take place, and if the government deals with other problems appropriately, the year 2008 will definitely herald the</p>	0,424

beginning of a new Nepal. We wish you a happy and prosperous New Year.

Paragraph 1: ktmpost/2008/01/02_year-of-hope_132912.txt The year that has just bade us 0.345
farewell was one of post-conflict chaos and confusion. Both the two dates announced for the constituent assembly elections came and went without the vote taking place. The past year will be remembered for violent movements, especially in the tarai, and a few but serious conflicts in the western tarai. Crime and political murders spread fast and hard in the tarai. The Young Communist League (YCL) formed last year also set new records in violence and extortion. The year also saw the height of anarchy, impunity and lawlessness. The streets of Kathmandu gained infamy for the most anarchic traffic in the whole world. This was the year Nepal declared itself a Federal Democratic Republic, but the decision needs to be endorsed which will probably happen in 2008. In a nutshell, the year 2007 was a year filled with turmoil.

Paragraph 2: Living in a country like Nepal, where 0.182
nothing is certain, it is very hard to be

ktmpost/2008/01/02_year-of-hope_132912.txt

optimistic. However, if we compare the present situation with the period before 2006, we have definitely come a long way. We have fewer people getting killed every day, and the most widely anticipated sustained peace looks achievable. 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. Interestingly, during the final days of 2007, the tarai movement seemed to turn peaceful, the YCL honored its supreme commander's promise and refrained from violent acts, and even the traffic police moved swiftly to restore traffic discipline. It seems we can hope much from the year 2008.

Paragraph 6:

NepaliTimes/2007/12/14305_peace_train.txt

Janajatis have also gone along saying 0.174 the elections are more important for now. Madhesi militant factions have denounced it, but the Forum will take part. The rightwing RJP and RPP dubbed the deal undemocratic. China, India and UNMIN cautiously welcomed the new agreement. The scepticism is understandable because by their past behaviors the seven parties give us no reason to trust them this time.

<p>Paragraph 8:</p> <p>NepaliTimes/2007/12/14305_peace_train.txt</p>	<p>The YCL has harmed the prospects of its own parent party in elections, but the risk is that their excesses will be used by the hard right to wreck elections. Together, the six-party alliance and Maoists must finally prove they are capable of governing.</p>	0.134
<p>Paragraph 7:</p> <p>NepaliTimes/2007/12/14305_peace_train.txt</p>	<p>Polls have been postponed twice. This time, if they fail to have elections by April, the leaders don't just lose face, they lose all legitimacy. Still, from now till April there are three spoilers. Monarchists will try their damndest to not allow elections to happen, an escalation in madesh violence would be a problem, and the wild followers of the Maoists will have to be defanged.</p>	0.106
<p>Paragraph 4:</p> <p>NepaliTimes/2007/12/14305_peace_train.txt</p>	<p>The mixed election system has been slightly modified as a face-saving device for the Maoists, but the essence of the earlier arithmetic remains in place. The 23-point agreement is what is already in the interim constitution and in previous accords.</p>	0.081
<p>Paragraph 2:</p> <p>NepaliTimes/2007/12/14305_peace_train.txt</p>	<p>So we have a 23-point agreement among the seven parties in which the communists commit</p>	0.069

	themselves, once again, to constituent assembly elections.
Paragraph 3: NepaliTimes/2007/12/14305_peace_train.txt	Nepal has been declared a republic, 0.057 but it will only take formal effect sometime in the middle of next year after it is ratified by the constituent assembly. But the king is in his palace, still paid a salary by taxpayers money.

5.4. Evaluation of the Retrieved Results

For evaluating the effectiveness of the query parameters and correspondingly the accuracy of the extracted texts in terms of relatedness with respect to the thesis statement, some test cases were developed and analysis of the output of the tool was conducted against the test cases. The test cases essentially represent the different compositions of the query parameters, which can include one or more combinations of the three different categories of query terms that have been defined earlier, i.e., exactly matching term, related term and domain-specific term. Similarly, the accuracy of the output or extracted texts in terms of relatedness is measured as a function of the total related texts divided by the total retrieved texts in each example. Then, mean accuracy for each test case is computed. An account of the test cases and their mean accuracies are provided in Table 5-7 and Table 5-8.

Table 5-7: Test cases and query parameters

Test cases	No.	of	Parameters	Examples	Accuracies (%)
Query expression 1 (QE_1)	2		<Exact term ₁ , Exact term ₂ >	<2007, violent>	17.86%
				<2007, chaos>	63.64%
				<2007, post-conflict>	100%
				<2007, confusion>	33.33%
Query expression 2 (QE_2)	3		<Exact term ₁ , Exact term ₂ , Exact term ₃ >	<2007, violent, year>	28.57%
				<2007, violent, chaos>	100%
				<2007, violent, confusion>	50%
Query expression 3 (QE_3)	3		<Exact term ₁ , Exact term ₂ , Similar term ₁ >	<2007, violent, crime>	16.66%

				<2007, post-conflict, impunity>	100%
				<2007, chaos, murders>	100%
				<2007, confusion, extortion>	100%
Query expression 4 (QE₄)	3	<Exact term ₁ , Exact term ₂ , Domain-specific term>		<2007, violent, Nepal>	100%
				<2007, chaos, CA>	100%
				<2007, post-conflict, YCL>	100%
Query expression 5 (QE₅)	4	<Exact term ₁ , Exact term ₂ , Exact term ₃ , Similar term ₁ >		<2007, year, chaos, anarchy>	100%
				<2007, violent, year, impunity>	50%
				<2007, violent, year, movements>	100%
Query expression 6 (QE₆)	4	<Exact term ₁ , Exact term ₂ , Exact term ₃ , Domain-specific term ₁ >		<2007, violent, year, Nepal>	100%

			<2007, violent, 100% year, CA elections>
			<2007, violent, 100% year, YCL>
Query expression 7 (QE₇)	5	<Exact term ₁ , Exact term ₂ , Exact term ₃ , Similar term ₁ , Domain- Specific term ₁ >	<2007, violent, 100% year, extortion, YCL>
			<2007, year, 100% confusion, movements, elections>
			<2007, year, chaos, 100% anarchy>

Table 5-8: Test cases and their mean accuracies

Test cases	Mean accuracies		
QE₁	54%	QE₂	60%
		QE₃	79%
		QE₄	100%
		QE₅	83%
		QE₆	100%
		QE₇	100%

The results in Table 5-8 are depicted in the form of a bar chart in Figure 5-5.

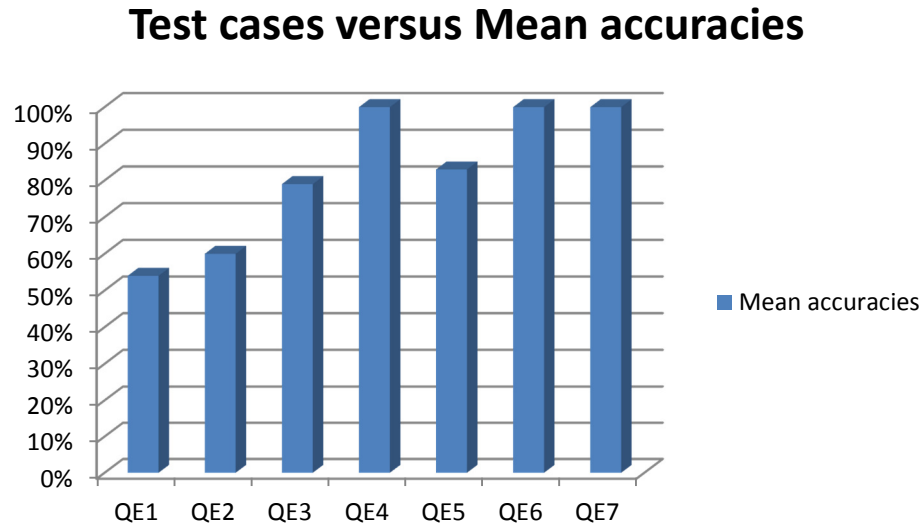


Figure 5-5: Bar chart diagram of test cases versus mean accuracies

As seen from Table 5-7, the test cases involve two to five parameters. QE_1 , for example, has two parameters whereas test cases ($QE_2 - QE_4$) have three parameters. Similarly, test cases QE_5 and QE_6 have four parameters. Finally, the test case QE_7 has five parameters. In general, the mean accuracy of the retrieval of the related texts is found to improve with the increase of the parameters in the query (See the mean accuracy scores for ($QE_1 - QE_7$)). However, it should be noted that even with the same number of parameters, the mean accuracy scores vary for the different test cases (See the mean accuracy scores for $QE_3 - QE_5$). The accuracy scores suggest that wherever the parameters ‘related term’ and ‘domain-specific term’ are involved in the query expression, the accuracy of the retrieval of the related texts is found to be higher compared to the ‘exact term’. This observation conveys a crucial message in adopting strategies for formulating efficient search queries for accurate text retrieval.

5.5. Extracting the Information of the Related Texts from the Annotated Texts

For the extraction of the opinion and argument information of the related texts from the annotated texts, the powerful text processing capabilities of the Grep tool and the annotations for the opinions and arguments in the opinionated texts corpus are exploited. It should be noted that the opinionated texts corpus is organized in such a manner that the root folder is the “Source” followed by “Year”, then after “Month” and then finally the text file. In that respect, the path of the text files follows the structure “Source/Year/Month/file_name.txt”. Here, “Source” refers to the publishing agency, for example, KTMPOST or BBC. Similarly, “Year” and “Month” respectively refer to the year and month to which the opinionated texts belong. Such an organization of file structure comes up with an added up advantage in the sense that the tool prints the complete path to the file thus providing useful information about the retrieved text like source and date together with all matching texts/paragraphs to the query with their location number, in this case the paragraph/argument ID. This information can be used for extracting the opinion and argument information of the related texts from the corresponding annotated texts. In Figure 5-6, a sample of the related texts among the retrieved texts returned by the tool for the thesis statement “2007 was a violent year full of chaos and confusion for Nepal” is presented. Similarly, the corresponding texts in the annotated form from multiple sources are provided in Figure 5-7 and Figure 5.8.

Source	Date	Title	Paragraph ID
<p>ktmpost/2008/01/02_year-of-hope_132912.txt:0: The year that has just bade us farewell was one of the post-conflict chaos and confusion. Both the two dates announced for the constituent assembly elections came and went without the voting taking place. The past year will be remembered for violent movements, especially in the tarai, and a few but serious conflicts in the western tarai. Crime and political murders spread fast and hard in the tarai. The Young Communist League (YCL) formed last year also set new records in violence and extortion. The year also saw the height of anarchy, impunity and lawlessness. The streets of Kathmandu gained infamy for the most anarchic traffic in the whole world. This was the year Nepal declared itself as a Federal Democratic Republic, but the decision needs to be endorsed, which will probably happen in 2008. In a nutshell, the year 2007 was a year filled with turmoil.</p> <p>NepaliTimes/2007/12/14305_peace_train.txt: 2: Nepal has been declared a republic, but it will only take a formal effect sometime in the middle of next year after it is ratified by the constituent assembly. But the king is in his palace, still paid a salary by taxpayers' money.</p>			

Figure 5-6: Sample of the related texts with respect to the thesis statement

```

<? xml version = "1.0" encoding = "utf-8">

<!DOCTYPE annotation.dtd>

<argumentation title = "Year of hope" date = "2008 Jan 02" url = http://ekantipur.com/the-kathmandu-post/2008/01/02\_editorial\_year-of-hope\_132912.html>"

<argument id = "0">

<statement id="0" type = "thesis statement" persuasion-effect= "yes" exp_type = "opinion" opinion_orientation = "negative"> The year that has just bade us farewell was one of the post-conflict chaos and confusion. </statement>

<statement id = "1" type= "support" support_type = "for" persuasion_effect= "yes"> Both the two dates announced for the constituent assembly came and went without the voting taking place.</statement>

<statement id = "2" type = "support" support_type = "for" persuasion_effect= "yes" exp_type= "opinion" opinion_orientation = "negative">The past year will be remembered for violent movements, especially in the tarai, and a few but serious conflicts in the western tarai.</statement>

<statement id = "3" type= "support" support_type= "for" persuasion_effect= "yes" exp_type= "opinion" opinion_orientation= "negative"> Crime and political murders spread fast and hard in the tarai.</statement>

<statement id = "4" type= "support" support_type= "for" persuasion_effect = "yes" exp_type = "opinion" opinion_orienatation = "negative"> The Young Communist League (YCL) formed last year also set new records in violence and extortion.</statement>

<statement id = "5" type = "support" support_type = "for" persuasion_effect = "yes" exp_type = "opinion" opinion_orientation = "negative"> The year also saw the height of anarchy, impunity and lawlessness. </statement>

```

Figure 5-7: Sample annotated text for the related text – Year of hope, KTMPOST

```

<? xml version = '1.0' encoding = 'UTF-8'?>

<!DOCTYPE annotation.dtd>

<argumentation title = "Peace train" date = "28 Dec 2007 – 03 Jan 2008" url = http://nepalitimes.com/issue/2007/12/28/Editorial/14305>

<argument id = "2">

<statement id = "0" type= "thesis statement" commitment_level= "high" exp_type= "opinion" opinion_orientation = "positive"> Nepal has been declared a republic, but it will only take a formal effect sometime in the middle of next year after it is ratified by the constituent assembly. </statement>

```

Figure 5-8: Sample annotated text for the related text – Peace train, Nepali Times

5.6. Constructing a Synthesis of Extracted Information from Annotated Texts

The task of constructing a synthesis of the extracted information from annotated texts involves utilizing the annotations in the form of attribute-values and organizing the thesis statement and related texts from different sources under some common attributes like “opinion_orientation”, “support_type” etc. along with the counts of supporting statements of a particular orientation. A sample of such a synthesis is provided in Figure 5-9 below.

<p>Global thesis statement: 2007 was a violent year full of post-conflict chaos and confusion.</p> <p>Source 1: Nepali Times, Date of publication: 28 Dec 2007 – 03 Jan 2008</p> <p>Source 2: The Kathmandu Post, Date of publication: Jan 02, 2008</p> <p>Thesis statement: Nepal has been declared a republic, but it will only take a formal effect sometime in the middle of next year after it is ratified by the constituent assembly. <The Kathmandu Post></p> <p>opinion_orientation: Positive (1)</p> <p>Support: For</p> <p>This was the year that Nepal declared itself as a Federal Democratic Republic, but the decision needs to be endorsed, which will probably happen in 2008. <The Kathmandu Post></p> <p>Support: Against</p> <p>But the king is in his palace, still paid a salary by taxpayers’ money. <Nepali Times></p> <p>opinion_orientation: Negative (8)</p> <p>Support: For</p> <p>The year that has just bade us farewell was one of the post-conflict chaos and confusion. <The Kathmandu Post></p> <p>The past year will be remembered for violent movements, especially in the tarai, and a few but serious conflicts in the western tarai. <The Kathmandu Post></p> <p>Crime and political murders spread fast and hard in the tarai. <The Kathmandu Post></p> <p>The Young Communist League (YCL) formed last year also set new records in violence and extortion. <The Kathmandu Post></p> <p>The year also saw the height of anarchy, impunity and lawlessness. <The Kathmandu Post></p>

Figure 5-9: Sample of synthesis of extracted texts with annotated attributes

A diagrammatic representation of the synthesis in Figure 5-9 is presented in Figure 5-10.

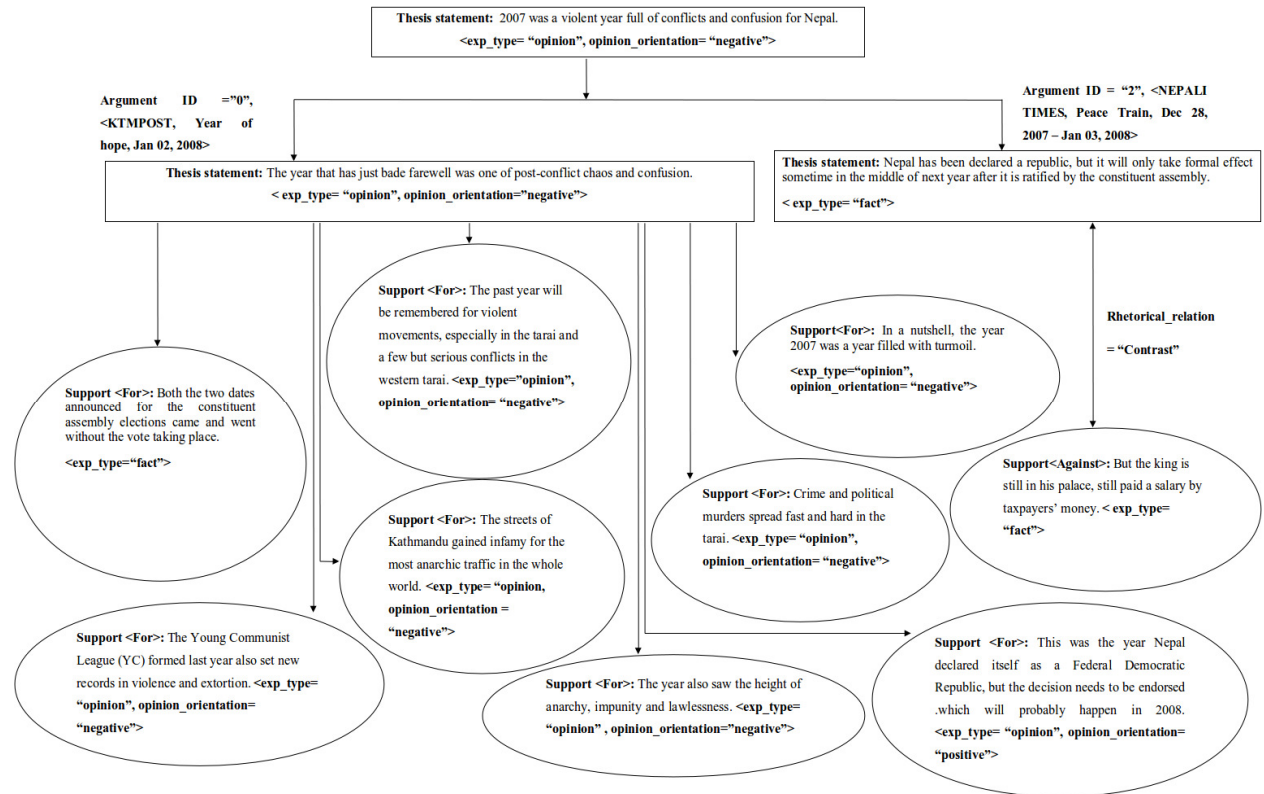


Figure 5-10 : Diagrammatic representation of the synthesis of extracted texts

As seen from Figure 5-10, the synthesized information represents itself as a graph structure. The structure has two different types of nodes – the thesis statement and the support statement. While a rectangular node represents the thesis statement, an oval node represents the support statement. The arcs between the nodes also are of two types – arcs with unidirectional arrow and arcs with bi-directional arrows. The unidirectional arcs point from the global thesis statement or the main topic or issue of discussion to the theses statements of the different arguments or one or more support statements from the thesis statement of an argument. The bi-directional arcs, on the other hand, refer to the rhetorical or coherence or discourse relations that can exist between the thesis statement and any one of the support statement or between two support statements. The arcs can be labeled for argument identity number as well as other useful information like source, date,

title etc. Similarly, the bi-directional arcs are labeled for different rhetorical relations. The nodes, on the other hand, not only has the statements or texts but also useful annotated information like “exp_type”, “opinion_orientation”, “Support”. As we see in Figure 5-10, from the global thesis statement, which essentially forms the basis for the search of related opinionated texts, the arcs branch to different arguments, which have their own set of thesis and support statements.

5.7. Conclusion

In this Chapter, we looked into the overall task of constructing a synthesis of opinions and arguments from multiple sources with respect to a common issue or topic of discussion. We started our discussion with the sub task of retrieving related texts with respect to the query terms based on the thesis statement, which represents the topic of discussion. We discussed the methodology of retrieving related texts and dealt with the experimental results and the overall accuracies. Then we briefly looked into the two processes, namely, the extraction of the annotation annotated information for the related texts and the construction of the synthesized information based on the extraction.

The whole process of constructing a synthesis of opinions and arguments from multiple sources with respect to a common issue or topic of discussion, as we saw in this Chapter, is dependent much on how reliably and accurately the annotations have been performed and also on whether the text retrieval tool returns the relevant and/or related texts or not. The performance of the text retrieval tool, on the other hand, is dependent on the formulation of the query, which involves the choice of appropriate query terms (exact terms, similar terms and domain-specific terms) and their combinations.

The next Chapter will discuss the Computational model for analyzing/synthesizing opinions and arguments from opinionated texts. The model represents a framework which puts the different sub-modules and the linguistic resources together responsible for producing the analysis of opinions and arguments in opinionated texts in the form of annotations. In addition to the analysis, the model also includes the opinion search and retrieval module and the synthesis module.

The two modules are responsible for finding related texts with respect to the query terms and finally extracting the respective annotations for the related texts in the form of attribute-values conveying important information about the opinions and arguments expressed in the texts.

CHAPTER 6 COMPUTATIONAL LINGUISTIC MODEL

In the earlier Chapters 3-5, we discussed several theories and algorithms for analyzing and synthesizing opinions and arguments in opinionated texts. Essentially, through our discussion in those chapters, we build the basis for the Computational Linguistic model. In this Chapter, we will see how the concepts and methods discussed earlier have been brought together within the implementation of the model as a demonstration of proof of concept.

6.1. The Components of the Model

The model comprises different components performing the following actions:

- 1) Fetch opinionated texts
- 2) Pre-process texts
- 3) Annotate texts
- 4) Retrieve related texts
- 5) Extract annotations for related texts
- 6) Produce a synthesis of opinions and arguments

The list above refers to the overall processes involved in the analysis and synthesis of opinionated texts in the implementation of the model. In **Error! Reference source not found.**, an overview of the high level architecture of the computational model is presented.

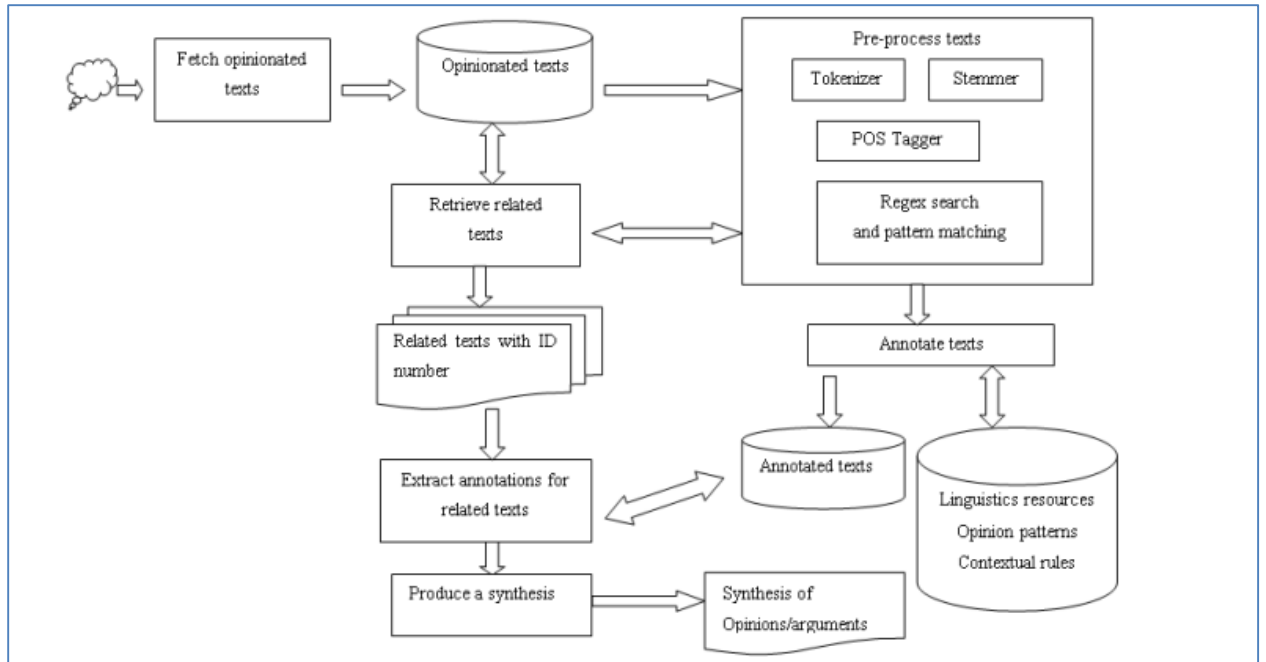


Figure 6-1: High level architecture of the computational model

In the next sections, a brief description of each of the processes is provided.

6.2. Fetch opinionated texts

The opinionated texts are fetched from a pre-selected set of online news portals. The time frame for downloading the opinionated texts has been set for 2007 to 2012 and the number of news portals being five, but these can be changed as per the needs. A self-executable script has been developed for the purpose of crawling. The opinionated texts are downloaded in both html file and plain text file formats. The downloaded texts are organized in a hierarchical fashion in terms of storing them in sub-folders within folders. For example, an editorial from the year “2012” and belonging to the “December” issue would be located inside the “12” sub-folder (month of December in numeral format) inside the “2012” folder. Such an organization greatly eases the subsequent accesses of the texts. With a view to capture the useful information about each editorial at the time of downloading, the script also saves certain meta data information like “Date”,

“Source”, “Author”, “URL”, “Gist” etc. in the form of comma separated values (CSV) file format for future use.

6.3. Pre-process texts

The selected opinionated texts undergo several standard forms of pre-processing like tokenization, stemming, POS tagging, regular expression search and pattern matching. For this purpose, standard libraries available in the Natural Language Toolkit (NLTK) package and the Grep tool are used. Pre-processing of texts is done for both retrieval of related texts and annotation of the texts.

6.4. Annotate texts

Based on the knowledge gathered from the linguistic resources (semantic tagset, discourse marker rules, sentiment/polarity lexicon, intensifier lexicon, reporting and modal verbs lexicon), opinion patterns and contextual rules, the automatic annotation tool annotates the texts from the plain texts files in extensible markup language (XML) file format. The annotated texts are stored in the annotated texts database for further use. The process of annotating texts has been described in Chapter 4.

6.5. Retrieve the related texts

Queries are passed to the database for related texts through the retrieval tool. The queries are terms (words or phrases) representing the thesis statement or issue of discussion formed by combining one or more logical operators AND, OR, NOT. The terms may be exact to the terms in the thesis statement, a related term, or a domain-specific term. The related texts to the query are individual texts/paragraphs from one or more opinionated texts sources with identity number (line or paragraph number) as well as meta data information like date and source.

6.6. Extract annotations for related texts

Using the information of the related texts, i.e., the text identity number and the meta data information of the source, the texts are then fetched from the annotated texts data bank together with their annotations for opinions and arguments. The annotated texts are essentially the thesis and support statements (opinions and arguments) in favor of or against the global thesis statement or issue of discussion.

6.7. Produce a synthesis of opinions and arguments

Finally, a synthesis of the related texts along with annotations for opinions and arguments is produced. The synthesis is in plain text format, with the related texts grouped under a common attribute-value like opinion orientation. The global thesis statement is placed at the top of the document followed by one or more local theses statements from multiple sources. For each of these local theses statements, one or more supports from different sources and belonging to different opinion orientations are presented.

The processes – retrieving related texts, extracting annotations for related texts and finally producing a synthesis of opinions and arguments have been described in Chapter 5.

6.8. Conclusion

In this Chapter, we looked into the overall Opinion and Argumentation Analysis/Synthesis system. We discussed the different components or modules of the overall system with a brief description of the processes involved in each of the modules. The next Chapter will discuss the contributions of this research and recommendations for future works.

CHAPTER 7 CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORKS

7.1. Contributions to Knowledge

The major contributions of the given work may be summarized as follows:

- 1) Synthesis of the concepts from a number of inter-related sub-disciplines of Artificial Intelligence and Computational Linguistics like argumentation theory, propositional logic and rhetorical analysis as well as opinion analysis in the lexical, phrase, sentence and paragraph level for addressing the task of opinion analysis and synthesis.
- 2) A good amount of resources for opinion and argumentation analysis has been developed. These resources include:
 - a. A semantic tagset for annotating argument and opinions in opinionated text. The tagset represents a collection of useful features that captures opinion analysis in the sentence level. In addition, it also has those features like “thesis statement”, “support statement”, “support type”, “rhetorical relations” required for analyzing opinions and arguments in the sentence and paragraph levels. The available annotation schemes only deal with annotating opinions and appraisal units as well as some rhetorical relations and discourse markers but lack the features of opinion analysis in terms of deeper linguistic analysis like persuasion effects and commitment level. Besides, the argumentation features required for analyzing and synthesizing arguments in opinionated texts are also lacking.
 - b. Manual annotation of 100 opinionated texts from both leading national and international news portals like the Nepali Times, The Kathmandu Post, BBC, Aljazeera and the Guardian – the major theme being “Socio-political”. The annotated texts can be used for multiple purposes – as a training data for the annotation tool for automatically annotating any new

texts and as a gold standard data for checking the accuracy of the automatically annotated output text.

- 3) Developed linguistic resources augmenting the existing available linguistic resources:
 - a. A Sentiment/Polarity lexicon with 300 positive and 800 negative terms specifically from the “socio-political” domain. Reasonably good results have been achieved together with this lexicon and available opinion lexicon in terms of automatic annotation of opinions in opinionated texts.
 - b. Similarly, some generalized opinion patterns helpful for analyzing opinions in the phrase and higher levels have been developed.
 - c. An intensifier and pre-modifier lexicon has been compiled from different sources. This lexicon is used in combination with some hand-crafted rules to determine the varying degrees of opinions. Currently, rules have been defined to capture at least seven degrees of opinion strength scales (-3 to 3, respectively, denoting “lowest”, “lower”, “low”, “average”, “high”, “higher”, “highest”)
 - d. A reporting verb and modal verb lexicon has been compiled from different sources. This lexicon is used for capturing the commitment or intent level of the opinions and statements. The consideration of reporting and modal verbs for determining the commitment or intent level adds an extra dimension for measuring the strength of opinions.
 - e. Discourse markers are crucial in signaling the presence of rhetorical relations in texts. By conducting an analysis of the collected opinionated texts corpus, some frequently occurring discourse markers have been identified and mapped to eight rhetorical relations useful for discourse analysis of opinionated texts. Further, a discourse markers and rhetorical relations lexicon has been compiled from different sources.
- 4) A feasibility study of the complete system for producing a synthesis of opinions and arguments on a common topic or issue of discussion from multiple sources has

been conducted via a computational linguistic model. The model suggests effective use of the developed linguistic resources and computational methods to achieve the goals of argumentation and opinion analysis and synthesis.

7.2. Extensions for Future works

The current work can be extended for further works in the future in the following fronts:

- 1) Replicating the current efforts for Nepali. Opinionated texts for Nepali from two online news portals – “ekantipur” and “nagariknews” have also been downloaded as part of the corpus. Preliminary study of the corpus has shown that the current framework of opinion and argument analysis and synthesis developed for English can be replicated for Nepali with considerable results.
- 2) Increasing the coverage of the current linguistic resources with more domain-specific opinion terms and expressions and contextual rules. A good amount of potentially opinion bearing terms and expressions (unigrams and bi-grams) has been gathered from the corpus (See Chapter 4). These, however, need to be further categorized as positive or negative. In addition to this, contextual rules to capture implicit discourse markers are also currently not considered in the system. The performance of the system is expected to greatly improve with the addition of these resources.
- 3) Incorporating Machine Learning capabilities to the automatic annotation tool. The current automatic annotation tool relies heavily on hand-crafted lexicon terms and rules to detect and annotate the opinion and argumentation aspects in opinionated texts. Adding Machine Learning capabilities to the tool can exploit the large corpus data that has been collected to further enhance the tagger in terms of automatic detection and tagging of opinion and argument terms and expressions.
- 4) Testing the linguistic resources for other domains. The linguistic resources developed have been tested basically in the socio-political domain. However, their reuse and applicability is yet to be tested in other domains.

- 5) Incorporating a Named Entity Recognition (NER) system for better handling facts and opinions. The current system distinguishes facts and opinions through some limited linguistic cues and rules. This can be better handled incorporating a NER system. Whether or not the existing NER systems for English are suitable for opinionated texts is another question for research.
- 6) The synthesis of opinions and arguments currently is quite basic in the sense that it assumes just the organization of annotated opinion and argumentation elements for the retrieved texts with respect to the global thesis statement. However, this can be further elaborated in terms of design in order to convey a more meaningful representation of the synthesis of opinions and arguments from multiple sources. Considering the role of entities with a primary focus could be one of the aspects to enhance the results of the retrieval of related texts. Future works in this direction could be towards integrating separate related texts in a more natural or fluent human-like expression to give a proper balanced account of the area.
- 7) The development of a fully automatic system for analysis and synthesis of the opinions and arguments from multiple sources and its evaluation – the current system has been tested for the proof of concept but it requires further work and evaluation in order to make it fully functional in real world settings.

References

Baccianella, S., Esuli, A., & Sebastiani, F. (2010). SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. *In Proceedings of the Languages Resource and Evaluation (LREC)*.

Benmara, F., Cesarano, C., Picariello, A., Reforgiato, D., & Subrahmanian, V. (2007). Sentiment Analysis: Adjectives and Adverbs are Better than Adjectives Alone. *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*.

Besnard, P., & Hunter, A. (2008). *Elements of Argumentation*. MIT Press.

Bolivar, A. (1994). The structure of newspaper editorials. In C. Malcom (Ed.), *Advances in written text analysis* (pp. 276-294). London: Routledge.

Bruce, R. F., & Wiebe, J. M. (2000). Recognizing subjectivity: a case study in manual tagging. *Natural Language Engineering* , 5 (2), 187-205.

Burton, G. O. (2014, 04 06). Retrieved 04 06, 2014, from "Silvia Rhetoricae" (rhetoric.byu.edu): <http://rhetoric.byu.edu/>

Capital Community College. (2014). *Adjectives*. Retrieved May 8, 2014, from Capital Community College - Hartford CT: <http://grammar.ccc.commnet.edu/grammar/adjectives.htm>

Capital Community College. (2014). *Adverbs*. Retrieved May 8, 2014, from Capital Community College - Hartford CT: grammar.ccc.commnet.edu/grammar/adverbs.htm

Capital Community College. (2014). *Helping and Modal Auxiliary Verbs*. Retrieved May 8, 2014, from Capital Community College - Hartford CT: <http://grammar.ccc.commnet.edu/grammar/auxiliary.htm>

Carlson, L., Marcu, D., & Okurowski, M. (2001). Building a Discourse-tagged Corpus in the Framework of Rhetorical Structure Theory. *In Proceedings of the Second Sigdial Workshop*

on Discourse and Dialogue. 16, pp. 1-10. Aalborg, Denmark: Annual Meeting of the ACL, Association for Computational Linguistics, Morristown, NJ.

Chesley, P., Vincent, B., Xu, L., & Srihari, R. (2006). Using Verbs and Adjectives to Automatically Classify Blog Sentiment. *In Proceedings of AAAI-CAAW-06, the Spring Symposium on Computational Approaches to Analyzing Weblogs*.

Cohen, J. (1960). A coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, XX (1).

Dainton, M., & Zelly, E. D. (2013). *Applying Communication Theory for Professional Life: A Practical Introduction* (3rd Edition ed.). Sage Publications, Inc.

Davdar, M., Hauff, C., & Jong de, F. (2011). Scope of Negation Detection in Sentiment Analysis. *In: Dutch-Belgian Information Retrieval Workshop, DIR 2011, 4 February 2011*. Amsterdam, the Netherlands.

Dowden, B. (1993). *Logical reasoning*. Belmont Calif: Wadsworth Publ. Comp.

Esuli, A., & Sebastiani, F. (2006). Determining term subjectivity and term orientation for opinion mining. *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics (EACL-2006)*.

Esuli, A., & Sebastiani, F. (2005). Determining the semantic orientation of terms through gloss classification. *In Proceedings of the 14th ACM international conference on Information and knowledge management* (pp. 617-624). New York, NY, USA: ACM.

Esuli, A., & Sebastiani, F. (2006). SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining. *In Proceedings of the 5'th Conference on Language Resources and Evaluation (LREC '06)*, (pp. 417-422).

Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics* (31), 931-952.

Government of South Australia. (2011). *Department for Education and Child Development(DECd)*. Retrieved June 26, 2013, from http://www.decd.sa.gov.au/accountability/files/links/link_156373.pdf

Hatzivassiloglou, V., & McKeown, K. R. (1997). Predicting the semantic orientation of adjectives. *In Proceedings of the Eighth Conference on European Chapter of the Association for Computational Linguistics (EACL '97)* (pp. 174-181). Stroudsburg, PA, USA: Association for Computational Linguistics.

Hatzivassiloglou, V., & Wiebe, J. M. (2000). Effects of Adjective Orientation and Gradability on Sentence Subjectivity. *In Proceedings of the 18'th Conference on Computational Linguistics (COLING '00). 1*, pp. 299-305. Stroudsburg, PA, USA: Association for Computational Linguistics.

Hatzivassiloglou, V., Klavans, J. L., Holcombe, M. L., Barzilay, R., Kan, M.-Y., & McKeown, K. (2001). Simfinder: A flexible clustering tool for summarization. *In Proceedings of the Workshop on Summarization in NAACL-01*.

Heerschop, B., Iterson, P. v., Hogenboom, A., Frasinicar, F., & Kaymak, U. (2011). Accounting for negation in sentiment analysis. *Proceedings of the 11th Dutch-Belgian information retrieval workshop : DIR 2011, February 4, 2011*, (pp. 38-39). Amsterdam, the Netherlands.

Hu, M., & Liu, B. (2004). Mining and Summarizing Customer Reviews. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004), Aug 22-25, 2004*. Seattle, Washington, USA.

J.J. Strossmayer University of Osijek. (2008). *DM-Discourse Markers*. Retrieved May 26, 2014, from Faculty of Economics in Osijek: http://www.efos.unios.hr/arhiva/index.php/component/docman/doc_details/190-dm-discourse-markers-170508

Kim, S.-M., & Hovy, E. (2006). Extracting Opinions, Opinion Holders, and Topics Expressed in Online News Media Text. *ACL Workshop on Sentiment and Subjectivity in Text*, Sydney, Australia.

Kolflaath, E. (2007). *The Strength of Arguments*. Department of Philosophy and Faculty of Law, University of Bergen.

Lin, D. (1998). Automatic retrieval and clustering of similar words. In *Proceedings of the 36'th Annual Meeting of the Association for Computational Linguistics and 17'th International Conference on Computational Linguistics (COLING-ACL-1998)*.

Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.

Liu, B., Hu, M., & Cheng, J. (2005). Opinion Observer: Analyzing and Comparing Opinions on the Web. *Proceedings of the 14'th International World Wide Web Conference (WWW-2005)*, May 10-14, 2005. Chiba, Japan.

London Metropolitan University. (2010). *Discourse Markers Sign-posts*. Retrieved May 8, 2014, from London Metropolitan University: <http://learning.londonmet.ac.uk/TLTC/connorj/WritingGroups/Writing/5%20discourse%20markers-signposts.pdf>

Mann, W., & Thompson, S. (1987). *Rhetorical structure theory: A theory of text organization*. Information Sciences Institute.

Marcu, D. (1998). A Surface-based Approach for Identifying Discourse Markers and Elementary Textual Units in Unrestricted Texts. In S. Manfred, L. Wanner, & E. Hovy (Ed.), *Proceedings of COLING-ACL Workshop on Discourse Relations and Discourse Markers*, (pp. 1-7). Montreal, Canada.

Marcus, M. P., Marcinkiewicz, M. A., & Santorini, B. (1993). Building a large annotated corpus of English: the penn treebank. *Computational Linguistics*, 19 (2), 313-330.

Martin, J., & White, P. R. (2005). *The Language of Evaluation: Appraisal in English*. London: Palgrave: Macmillan.

Nepali Times. (17 - 23 July, 2009). Audacity to Hope. *Editorial* . Kathmandu, Bagmati, Nepal: Nepali Times .

O'Keefe, D. J. (2002). *Persuasion: Theory and Research (Current Communication: An Advanced Text)*. SAGE Publications, Inc; 2nd edition.

Palau, R. M., & Moens, M.-F. (2009). Argumentation Mining: The Detection, Classification and Structure of Arguments in Text. *The 12th International Conference on Artificial Intelligence and Law, June 8-12, 2009* (pp. 98-107). Barcelona, Spain: ACM.

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up?: sentiment classification using machine learning techniques. *ACL-02 conference on Empirical methods in natural language processing*, (pp. 79-86).

Polyani, L. (2003). *The Linguistic Structure of Discourse. The Handbook of Discourse Analysis*. (S. E.D., Ed.)

Power, R., Doran, C., & Scott, D. (2000). Generating Embedded Discourse Markers from Rhetorical Structure. In *Proceedings of the 7th European Workshop on Natural Language Generation (EWNLG'99)*.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. New York: Longman.

Read, J., Hope, D., & Carroll, J. (2007). Annotating Expressions of Appraisal in English. *Proceedings of the ACL 2007 Linguistic Annotation Workshop*. Prague, Czech Republic.

Resnik, P., & Lin, J. (2010). Evaluation of NLP Systems. In A. Clark, C. Fox, & S. Lappin, *The Handbook of Computational Linguistics and Natural Language Processing*. Wiley-Blackwell.

Riloff, E. (1996). Automatically generating extraction patterns from untagged text. *In Proceedings of the AAAI-1996.*

Riloff, E., & Wiebe, J. (2003). Learning Extraction Patterns for Subjective Expressions. *In Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing (EMNLP '03).* (pp. 105-112). Stroudsburg, PA, USA: Association for Computational Linguistics.

Riloff, E., Patwardhan, S., & Wiebe, J. (2006). Feature subsumption for opinion analysis. *In Proceedings of the Conference on Empirical Methods in Natural Language Processing(EMNLP-2006).*

Riloff, E., Wiebe, J., & Wilson, T. (2003). Learning Subjective Nouns Using Extraction Pattern Bootstrapping. *In Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003-Volume 4 (CONLL '03).* 4, pp. 25-32. Stroudsburg, PA, USA: Association for Computational Linguistics.

Roberts-Miller, T. (2008). *Understanding Misunderstanding: How to do a Rhetoric Analysis*. Retrieved July 11, 2013, from <http://www.drw.utexas.edu/roberts-miller:>
<http://www.drw.utexas.edu/roberts-miller/handouts/rhetorical-analysis>

Rosen, K. H. (1998). *Discrete Mathematics and its Applications*. William C Brown Pub; 4th Edition.

Salton, G., & Lesk, M. E. (1968). Computer Evaluation of Indexing and Text Processing. *Journal of the ACM (JACM)* , 15 (1), 8-36.

Sharpling, G. (2012, August 22). *Discourse Markers: The University of Warwick*. Retrieved May 26, 2014, from http://www2.warwick.ac.uk/fac/soc/al/learning_english/leap/grammar/discourse/

Sherif, C. W., Sherif, M., & Nebergall, R. E. (1965). *Attitude and Attitude Change: The Social judgement involvement Approach*. Greenwood Publishing Group, Inc.

Sherif, M., & Hovland, C. (1961). *Social Judgement: Assimilation and contrast effects in communication and attitude change*. Yale University Press.

Simons, H. W. (1976). *Persuasion: Understanding, Practice, and Analysis*. Addison-Wesley Longman, Inc.

Sporleder, C., & Lascarides, A. (2008). Using Automatically Labelled Examples to Classify Rhetorical Relations: An Assessment. *Natural Language Engineering* , 14 (3), 369-416.

Stoyanov, V., Cardie, C., Litman, D., & Wiebe, J. (2004). Evaluating an Opinion Annotation Scheme Using a New Multi-Perspective Question and Answer Corpus. (Q. S. Wiebe, Ed.) *Computing Attitude and Affect in Text: Theory and Practice* , 77-89.

Sweneey, C. (2010, December 07). *Improving Writing | Discourse Markers : A Teacher's Guide and Toolkit*. Retrieved May 26, 2014, from EnglishEdu - English Teaching & Learning Resources: <http://english.edusites.co.uk/article/improving-writing-discourse-markers-a-teachers-guide-and-toolkit/>

Taboada, M. (2009). Implicit and Explicit Coherence Relations. (J. Rankema, Ed.) *Discourse, of Course* , 127-140.

Tabodoe, M., & Renkema, J. (2008). Discourse Relations Reference Corpus.

The Kathmandu Post. (2008, January 02). Year of hope. Kathmandu, Bagmati, Nepal: The Kathmandu Post.

The University of Adelaide. (2009). *Reporting Verbs - Learning Guide*. Retrieved May 8, 2014, from The University of Adelaide: http://www.adelaide.edu.au/writingcentre/learning_guides/learningGuide_reportingVerbs.pdf

Toulmin, S. (1958). *The uses of argument*. Cambridge University Press, Cambridge.

Turney, P. D. (2002). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. *In Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL '02)* (pp. 417-424). Stroudsburg, PA, USA: Association for Computational Linguistics.

Van Emeran, F. H., & Grootendorst, R. (2004). *A Systematic Theory of Argumentation. The pragma-dialectical approach*. Cambridge University Press.

Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation Schemes*. Cambridge University Press.

Wiebe, J. M. (2000). Learning Subjective Adjectives from Corpora. *In Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence* (pp. 735-740). AAAI Press.

Wiebe, J. M., & Riloff, E. (2005). Creating subjective and objective sentence classifiers from unannotated texts. *In Proceedings of the Computational Linguistics and Intelligent Text Processing*, (pp. 486-497).

Wiebe, J. M., Bruce, R. F., & O'Hara, T. P. (1999). Development and use of a gold-standard data set for subjectivity classifications. *In Proceedings of the 37'th Annual Meeting of the Association for Computational Linguistics on Computational Linguistics (ACL '99)* (pp. 246-253). Stroudsburgh, PA, USA: Association for Computational Linguistics.

Wiebe, J. M., Wilson, T., Bruce, R. F., Bell, M., & Martin, M. (2004). Learning subjective language. *Computational Linguistics* , 30 (3), 277-308.

Wiebe, J., Wilson, T., & Cardie, C. (2005). Annotating Expressions of Opinions and Emotions in Language. *Language Resources and Evaluation* , 1 (2).

Wiegand, M., Balahur, A., Roth, B., Klakow, D., & Montoyo, A. (2010). A survey on the role of negation in sentiment analysis. *In Proceedings of the Workshop on Negation and*

Speculation in Natural Language Processing (NeSp-NLP '10), Roser Morante and Caroline Sporleder (Eds.) (pp. 60-68). Stroudsburg, PA, USA: Association for Computational Linguistics.

Wilson T., W. J. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. *Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, (pp. 347-354).

Wilson, T. A. (2008). *Fine-grained subjectivity and sentiment analysis: Recognizing the intensity, polarity, and attitudes of private states* . PhD Thesis, University of Pittsburgh, Department of Computer Science.

Wilson, T., & Wiebe, J. (2005). Annotating Attributions and Private States. *In Proceedings of the ACL 2005 Workshop: Frontiers in Corpus Annotation II: Pie in the Sky*, (pp. 53-60).

Wilson, T., & Wiebe, J. (2003). Annotating Opinions in World Press. *Proceedings of the SIGdial-03*, (pp. 13-22).

Wilson, T., Wiebe, J., & Hwa, R. (2006). Recognizing Strong and Weak Opinion Clauses. *Computational Intelligence* , 73-99.

Yu, H., & Hatzivassiloglou, V. (2003). Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. *In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2003)*.

Zhang, L., & Liu, B. (2011). Identifying noun product features that imply opinions. *In Proceedings of the Annual Meeting of the Association for Computational Linguistics (short paper) ACL-2011*. Association for Computational Linguistics.

Zhang, W., & Yu, C. (2007). *UIC at TREC 2007 Blog Report*.

Zhang, W., Yu, C., & Meng, W. (2007). Opinion retrieval from blogs. *In Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management (CIKM '07)*. New York, NY, USA: ACM.

Bibliography

Abdalla, R., & Teufel, S. (2006). A Bootstrapping Approach to Unsupervised Detection of Cue Phrase Variants. *Proceedings of the 21'st International Conference on Computational Linguistics and 44'th Annual Meeting of the ACL* (pp. 921-928). Sydney: Association of Computational Linguistics.

Allen, J. (2003). *Natural Language Understanding* (2nd Edition ed.). Delhi: Pearson Education (Singapore) Pte. Ltd, Indian Branch, 482 F.I.E. Patparganj, Delhi 110 092, India.

Asr, F. T., & Demberg, V. (2012). Measuring the Strength of Linguistic Cues for Discourse Relations. *24'th International Conference on Computational Linguistics. Proceedings of the Workshop on Advances in Discourse Analysis and its Computational Aspects (ADACA)* (pp. 33-41). COLING 2012.

Benmara, F., Cesarano, C., Picariello, A., Reforgiato, D., & Subrahmanian, V. (2007). Sentiment Analysis: Adjectives and Adverbs are Better than Adjectives Alone. *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*.

Bird, S., Klein, E., & Loper, E. (2009). *NLTK Book: Natural Language Processing with Python*. O'Reilly.

Blanco, E., & Moldovan, D. (2011). Some Issues on Detecting Negation from text. *Twenty-Fourth International FLAIRS Conference*. AAAI Publications.

Budanitsky, A., & Hirst, G. (2006). Evaluating WordNet-based Measures of Lexical Semantic Relatedness. *Computational Linguistics* , 32 (1), 13-47.

Chaudhari, D. L., Damani, O. P., & Laxman, S. (2011). Lexical co-occurrence, Statistical Significance and Word Association. *In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing* (pp. 1058-1068). Edinburgh, Scotland, UK: Association for Computational Linguistics.

Chesley, P., Vincent, B., Xu, L., & Srihari, R. (2006). Using Verbs and Adjectives to Automatically Classify Blog Sentiment. *In Proceedings of AAAI-CAAW-06, the Spring Symposia on Computational Approaches to Analyzing Weblogs.*

Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics* , 16 (1), 22-29.

Cohen, R. (1987). Analyzing the Structure of Argumentation Discourse. *Computational Linguistics* , 11-24.

Corley, C., & Mihalcea, R. (2005). Measuring the Similarity of Texts. *In Proceedings of the ACL Workshop on Empirical Modeling of Semantic Equivalence and Entailment (EMSEE '05)* (pp. 13-18). Stroudsburg, PA, USA: Association for Computational Linguistics.

Corston-Oliver, S. (1998). *Identifying the Linguistic Correlates of Rhetorical Relations.* TechReport, Microsoft Research, MSR-TR-98-73.

Cramer, I. (2008). How well do semantic relatedness measures perform?: a meta-study. *In Proceedings of the 2008 Conference on Semantics in Text Processing (STEP '08)* (pp. 59-70). Stroudsburg, PA, USA: Association for Computational Linguistics.

Dagan, I., & Glickman, O. (2004). Probabilistic Textual Entailment: Generic Applied Modeling of Language Variability. *Learning Methods for Text Understanding and Mining Workshop.*

Dagan, I., Dolan, B., Magnini, B., & Roth, D. (2009). Recognizing textual entailment: Rational, evaluation approaches. *Natural Language Engineering* , XV (4).

Dagan, I., Shaul, M., & Shaul, M. (1995). Contextual Word Similarity and Estimation from Sparse Data. *Computer, Speech and Language* , 9, 123-152.

duVerle, D. A., & Predinger, H. (2009). A Novel Discourse Parser Based on Support Vector Machine Classification. *In Proceedings of the Joint Conference of the 47'th Annual Meeting of the ACL. Joint Conference on Natural Language Processing of the AFNLP*. 2, pp. 665-673. Stroudsburg, PA, USA: Association for Computational Linguistics.

Friedl, J. E. (2006). *Mastering Regular Expressions* (3rd Edition ed.). O'Reilly Media, Inc.

Gabrilovich, E., & Markovitch, S. (2007). Computing Semantic Relatedness Using Wikipedia-based Explicit Semantic Analysis. *In Proceedings of the 20'th International Joint Conference on Artificial Intelligence (IJCAI '07)* (pp. 1606-1611). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Ghosh, S., Riccardi, G., & Johansson, R. (2012). Global Features for Shallow Discourse Parsing. *In Proceedings of the 13'th Annual Meeting of the Special Interest Group on Discourse and Dialog (SIGDIAL '12)* (pp. 150-159). Stroudsburg, PA, USA: Association for Computational Linguistics.

Grosz, B., & Sidner, C. (1986). Attention, Intentions and the Structure of Discourse. *Computational Linguistics* , 175-204.

Hirschberg, J., & Litman, D. (1993). Empirical Studies on the Disambiguation of Cue Phrases. *Association for Computational Linguistics* , 19 (3).

Hyland, K. (1998). Persuasion and Context: The Pragmatics of Academic Metadiscourse. *Journal of Pragmatics* , 30 (4), 437-455.

Islam, A., Milios, E., & Keselj, V. (2012). Comparing Word Relatedness Measures Based on Google n-grams. *In Proceedings of COLING 2012:posters* (pp. 495-506). Mumbai: COLING 2012.

Iwanska, L. M., & Shapiro, S. C. (Eds.). (2008). *Natural Language Processing and Knowledge Representation, Language for Knowledge and Knowledge for Language*. Hyderabad: University Press (India) Private Limited.

Jia, L., Yu, C., & Meng, W. (2009). The effect of negation on sentiment analysis and retrieval effectiveness. *In Proceedings of the 18th ACM conference on Information and Knowledge Management (CIKM '09)* (pp. 1827-1830). New York, NY, USA: ACM.

Jiang, J., & Conrath, D. (1997). Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy. *In Proceedings of the International Conference on Research in Computational Linguistics*, (pp. 19-33).

Johansson, R., & Moschitti, A. (2010). Syntactic and Semantic Structure for Opinion Expression Detection. *In Proceedings of the Fourteenth Conference on Computational Natural Language Learning*, (pp. 67-76).

Jones, K. S. (1988). *A Statistical Interpretation of Term Specificity and its Application in Retrieval* (Taylor Graham Series in Foundations of Information Science ed., Vol. 3). (P. Willett, Ed.) London, UK: Taylor Graham Publishing London, UK.

Jurafsky, D., & Martin, J. H. (2009). *Speech and Language Processing, An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (2nd Edition ed.). New Jersey: Pearson Education, Inc.. Upper Saddle River, New Jersey 07458.

Kirschner, P., Buckingham-Shum, S., & Carr, C. (Eds.). (2003). *Visualizing Argumentation - Software Tools for Collaborative and Educational Sense-Making*. London.

Knott, A., & Dale, R. (1994). Using Linguistic Phenomena to Motivate a Set of Coherence Relations. *Discourse Processes*, 18 (1), 35-62.

Knott, A., & Sanden, T. (1998). The Classification of Coherence Relations and their Linguistic Markers: An Exploration of Two Languages. *Journal of Pragmatics*, 30 (2), 135-175.

Kolb, P. (2009). Experiments on the Difference Between Semantic Similarity and Relatedness. *Proceedings of the 17'th Nordic Conference of Computational Linguistics NODALIDA* (pp. 81-88). Northern European Association for Language Technology.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An Introduction to Latent Semantic Analysis. *Discourse Processes* , 25, 259-284.

Le Thanh, H., Abeysinghe, G., & Huyck, C. (2004). Automated Discourse Segmentation by Syntactic Information and Cue Phrases. *Proceedings of Artificial Intelligence and Applications (AIA2004)*.

Leacock, C., Miller, G. A., & Chodorow, M. (1998). Using Corpus Statistics and WordNet Relations for Sense Identification. *Computational Linguistics* , 24 (1), 147-165.

Lesk, M. E. (1986). Automatic Sense Disambiguation Using Machine Readable Dictionaries: How to Tell a Pine from an Ice Cream Cone. *In Proceedings of the SIGDOC Conference*. Toronto.

Levin, B. (1993). *English Verb Classes and Alternations - A Preliminary Investigation*. Chicago and London: The University of Chicago Press.

Li, Y., Bandar, Z. A., & McLean, D. (2003). An approach for measuring Semantic Similarity between words Using Multiple Information Sources. *IEEE Transactions on Knowledge and Data Engineering* , XV (4).

Lin, C. Y., & Hovy, E. H. (2003). Automatic Evaluation of Summaries Using N-grams Co-occurrence Statistics. *In Proceedings of the Human Language Technology Conference (HLT-NAACL 2003)*. Edmonton, Canada.

Lintean, M., & Rus, V. (2012). Measuring Semantic Similarity in Short Texts through Greedy Pairing and Word Semantics. *In Proceedings of the 25'th International Florida Artificial*

Intelligence Research Society Conference. Association for the Advancement of Artificial Intelligence (www.aaai.org).

Liu, B. (2011). *Web Data Mining - Exploring Hyperlinks, Contents, and Usage Data* (2nd Edition ed.). Springer.

Liu, S., Liu, F., Yu, C., & Meng, W. (2004). An Effective Approach to Document Retrieval via Utilizing WordNet and Recognizing Phrases. In *Proceedings of the 27'th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '04)* (pp. 266-272). New York, NY, USA: ACM.

Marcu, D. (1998). A Surface-based Approach for Identifying Discourse Markers and Elementary Textual Units in Unrestricted Texts. In S. Manfred, L. Wanner, & E. Hovy (Ed.), *Proceedings of COLING-ACL Workshop on Discourse Relations and Discourse Markers*, (pp. 1-7). Montreal, Canada.

Mihalcea, R., Corley, C., & Strapparava, C. (2006). Corpus-based and Knowledge-based Measures of Text Semantic Similarity. American Association for Artificial Intelligence(www.aaai.org).

Mitchel, T. (1997). *Machine Learning*. McGraw Hill.

Oates, B. J. (2005). *Researching Information Systems and Computing*. SAGE Publications Ltd.

Palau, R. M., & Moens, M.-F. (2009). Argumentation Mining: The Detection, Classification and Structure of Arguments in Text. *The 12'th International Conference on Artificial Intelligence and Law, June 8-12, 2009* (pp. 98-107). Barcelona, Spain: ACM.

Pappas, N., Katsimpras, G., & Stamatatos, E. (2013). Distinguishing the Popularity Between Topics: A System for Up-to-date Opinion Retrieval and Mining in the Web. *In*

Proceedings of the 14th International Conference on Intelligent Text Processing and Computational Linguistics. Samos, Greece: ACM.

Pilehvar, M. T., Jurgens, D., & Navigli, R. (2013). Align, Disambiguate and Walk: A Unified Approach for Measuring Semantic Similarities. *In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (pp. 1341-1351). Sofia, Bulgaria: Association for Computational Linguistics.

Pirro, G., & Euzenat, J. (2010). A feature and information theoretic framework for semantic similarity and relatedness. In P. F. Patel-Schneider, Y. Pan, P. Hitzler, P. Mika, & L. Zhang (Ed.), *In Proceedings of the 9th International semantic web conference on the Semantic Web - Volume Part I (ISWC '10)* (pp. 615-630). Berlin, Heidelberg: Springer-Verlag.

Pitler Emily, N. A. (2008). *Easily Identifiable Discourse Relations*. Department of Computer & Information Science, Technical Reports (CIS), University of Pennsylvania.

Resnik, P. (1995). Using Information Content to Evaluate Semantic Similarity in a Taxonomy. In C. S. Mellish (Ed.), *In Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI '95)*. I, pp. 448-453. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Rocchio, J. (1971). *Relevance Feedback in Information Retrieval*. Englewood Cliffs, New Jersey: Prentice Hall, Inc.

Rottenberg, A.T., & Winchell, D.H. (2011). *The Structure of Argument* (7th Edition ed.). Bedford/St. Martin's.

Rottenberg, A.T., & Winchell, D.H. (2012). *Elements of Argument - A Text and Reader* (10th Edition ed.). Bedford/St. Martin's.

Russel, S., & Norvig, P. (2009). *Artificial Intelligence: A Modern Approach* (3rd Edition ed.). Prentice Hall.

Salahli, M. A. (2009). An Approach for Measuring Semantic Relatedness Between Words via Related Terms. *Mathematical and Computational Applications* (1), 55-63.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Process Management* , 24 (5), 513-523.

Salton, G., Singhal, A., Mitra, M., & Buckley, C. (1976). Automatic Text Structuring and Summarization. *Information Processing and Management* , 2 (32).

Saric, F., Glavas, G., Karan, M., Snajdar, J., & Basic, B. D. (2012). TakeLab:Systems for Measuring Semantic Text Similarity. *First Joint Conference on Lexical and Computational Semantics* (pp. 441-448). Montreal, Canada: Association for Computational Linguistics.

Sebesta, R. W. (2008). *Concepts of Programming Languages* (8th Edition ed.). Pearson Education, Inc.

Sethi, R. (1996). *Programming Languages: concepts & constructs* (2nd Edition ed.). Addison-Wesley Publishing Company.

Shuttleworth, J. (1999). *Editorial Writing (Living Language Series)*. (G. Keith, & J. Shuttleworth, Eds.) Hodder and Stoughton Ltd.

Somasundaran, S., Namata, G., Wiebe, J., & Getoor, L. (2009). Supervised and Unsupervised Methods in Employing Discourse Relations for Improving Opinion Polarity Classification. *In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, (pp. 170-179). Singapore.

Soricut, R., & Marcu, D. (2003). Sentence Level Discourse Parsing Using Syntactic and Lexical Information. *Proceedings of the HLT-NAACL 2003 Main Papers*, (pp. 149-156). Edmonton.

Sporleder C., L. A. (2005). Exploiting Linguistic Clues to Classify Rhetorical Relations. In *Proceedings of Recent Advances in Natural Language Processing (RANLP)*. Bulgaria.

Sporleder, C., & Lascarides, A. (2008). Using Automatically Labelled Examples to Classify Rhetorical Relations: An Assessment. *Natural Language Engineering* , 14 (3), 369-416.

Terra, E., & Clarke, C. (2003). Frequency Estimates for Statistical Word Similarity Measures. In *Proceedings of HLT-NAACL, Main Papers*, (pp. 165-172). Edmonton.

Teufel S. (1999). *Argumentative Zoning: Information Extraction from Scientific Text*. PhD Thesis, University of Edinburgh.

Teufel, S. (1998). Meta-discourse Markers and Problem-Structuring in Scientific Articles. In *Proceedings of the ACL-98 Workshop on Discourse Structure and Discourse Markers*.

Tsatsaronis, G., Varlamis, I., & Vazirgiannis, M. (2010). Text Relatedness Based on Word Thesaurus. *Journal of Artificial Intelligence Research* , 37 (1), 1-40.

Turney, P. D. (2001). Mining the Web for Synonyms: PMI-IR versus LSA on TOEFL. In *Proceedings of the 12'th European Conference on Machine Learning (EMCL '01)* (pp. 491-502). London, UK: Springer-Verlag.

Walton, D. N. (2002). *Legal Argumentation and Evidence*. Penn State Press, 2002.

Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation Schemes*. Cambridge University Press.

Whitelaw, C., Garg, N., & Argamon, S. (2005). Using appraisal groups for sentiment analysis. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management (CIKM '05)* (pp. 625-631). New York, NY, USA: ACM.

Wu, Z., & Palmer, M. (1994). Verbs Semantics and Lexical Selection. *In Proceedings of the 32'nd Annual Meeting of Association for Computational Linguistics (ACL '94)*. (pp. 133-138). Stroudsburg, PA, USA: Association for Computational Linguistics.

Yeh, E., Ramage, D., Manning, C. D., Agirre, E., & Soroa, A. (2009). WikiWalk: random walks on Wikipedia for semantic relatedness. *In Proceedings of the 2009 Workshop on Graph-based Methods for Natural Language Processing (TextGraphs-4)* (pp. 41-49). Stroudsburg, PA, USA: Association for Computational Linguistics.

Zhang, W., Shuang, L., Yu, C., Sun, C., Liu, F., & Meng, W. (2007). Recognition and Classification of Noun Phrases in Queries for Effective Retrieval. *In Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management (CIKM '07)* (pp. 711-720). New York, NY, USA: ACM.

Appendix A Publications

Bal, B. K. (2014). Analyzing Opinions and Argumentation in News Editorials and Op-Eds. *International Journal of Advanced Computer Science and Applications - IJACSA, Special Issue on Natural Language Processing , IV* (1).

Bal, B. K., & Saint-Dizier, P. (2010). Towards Building Annotated Resources for Analyzing Opinions and Argumentation in News Editorials. *LREC*. Malta: ELRA.

Bal, B. K. (2009). Towards an Analysis of Opinions in News Editorials: How Positive was the year? *In Proceedings of the 8th International Conference on Computational Semantics*. Tilburg.

Bal, B. K., & Saint-Dizier, P. (2009). Have we had Hard Times or Cosy Times? A Discourse Analysis of Opinions Expressed over Socio-political Events in News Editorials. *In Proceedings of the 7th International Conference on Natural Language Processing (ICON 2009)*. Hyderabad, India: Macmillan Publishers India Ltd.

Bal, B. K., & Saint-Dizier, P. (2009). Towards an Analysis of Argumentation Structure and Strength of Arguments in News Editorials. *AISB Symposium on Persuasive Technologies* . Edinburgh.

Bal, B. K., & Saint-Dizier, P. (2009). Who speaks for whom? Towards analyzing opinions in news editorials. *In Proceedings of the Eighth International Symposium on Natural Language Processing*. Bangkok: SNLP.

Appendix B Sample of editorials and blogs from the corpus

TITLE@Maoists' double standard

DATE@2007 May 05

URL@<http://ekantipur.com/the-kathmandu-post/2007/05/05/editorial/maoists-double-standard/108572.html>

MAY 04 - A report of the UN Office of the High Commissioner for Human Rights in Nepal (OHCHR-Nepal), issued last week, manifests the glaring facts about the CPN-Maoist. In the report the OHCHR-Nepal has starkly said that the Maoist cadres aren't complying with their party's commitments and are not respecting the rights of the Internally Displaced Persons (IDPs) to voluntarily and safely return home. The UN rights body has also expressed a serious concern over the continued existence of the Maoists' parallel government and their Kangaroo courts. The scrutiny by the UN rights body conspicuously underlines a strong message that the Maoists haven't abandoned the culture of intimidation. The former rebels' non-committal attitude towards their promises seems to be propelled by their instantaneous fear that the people, who suffered in the hands of the Maoists, will claim the price of their pain once the rebels lose their control. Equally fair observation is that the Maoists want to continue political vacuum in the villages. For them, the return of the IDPs means the return of other political parties in the villages. But the Maoists don't have the right to obstruct political activities in villages. Ironically, the Maoist leadership keeps on reiterating its promise that it will allow the return of the IDPs. Hence it encapsulates the Maoists' carrot and stick approach.

The return of the IDPs and their property was a cutting issue when the interim government was being formed in March. The prime minister and all the political parties had demanded a sincere commitment from the Maoist leadership before including it in the government. The Maoists had agreed in forming a committee of all political parties and immediately facilitating the return of the

IDPs and their property, but they didn't abide by their commitment and didn't act accordingly. The government statistics state that more than 100,000 people in Nepal were displaced during the decade-long insurgency, though it hasn't been independently verified. Although the Maoist leadership had announced that it would dissolve its parallel government and Kangaroo courts across the country, they have continued. Last week alone, police engaged in a brief scuffle with the Maoist cadres in Kathmandu in an attempt to seize the Maoists' vehicle that had the number plate of theso-called Maoist Third Division-4 of PLA. Fortunately, officials of the UN Mission in Nepal (UNMIN) took the vehicle and stopped the possible violence.

Since the Maoists have joined the interim government, they also have the responsibility to enforce law and order in the country. The Maoist leadership can't play to the gallery. It raises a serious question over the Maoists' commitment on democracy and human rights. The Post believes that the Maoists' double standard is weakening the eight-party alliance. The seven political parties in the current coalition government are growing more suspicious of the Maoists just because of the huge gap in their words and practice. The Post calls on the Maoist leadership to immediately initiate the process of managing a dignified and safe return of the IDPs and respect its commitment to democracy and human rights.

TITLE@Egypt and the Palestinian question

AUTHOR@Abdullah Al-Arian

DATE@05 Feb 2011

#URL@<http://www.aljazeera.com/indepth/opinion/2011/02/20112512042978473.html>

Along with the laundry list of domestic grievances expressed by Egyptian protesters calling for an end to the regime of Hosni Mubarak, the popular perception of Egypt's foreign policy has also been a focal point of the demonstrations.

Signs and chants have called on Mubarak to seek refuge in Tel Aviv, while his hastily appointed vice-president, Omar Suleiman, has been disparaged as a puppet of the US. Egypt's widely publicised sale of natural gas to Israel at rock bottom prices has featured in many refrains emanating from the crowds.

The widespread view among Egyptians that the regime has served the interests of the West has not been helped by Israel's call for world leaders to support Mubarak, or the apparent unwillingness by American officials to give the protests their full backing.

Plummeting status

In the shadow of the current cries to topple the Egyptian regime, the Mubarak government has had a tough time keeping its role in international affairs out of public view.

In the area where Egypt's foreign policy apparatus has served US interests most directly, Israel's security, the Mubarak regime's complicity in the failure to establish a Palestinian state has become widely publicised in recent years. Its role in placing the stranglehold on the people of Gaza, in conjunction with Israel, has seen Egypt's status in the region plummet to a level it has not reached in decades.

The Palestine Papers, the leaked internal documents of the Palestinian Authority (PA) that were recently exposed by Al Jazeera, provide further confirmation of Egypt's role in the impasse between Israeli and Palestinian negotiators.

While much of the coverage of the Palestine Papers has focused on the unprecedented concessions offered by Palestinian negotiators, and how swiftly they were spurned by Israeli and American representatives, Egypt's role as an instrument for added pressure stands out from the internal records.

As the peace process broke down over the past decade, Egypt was a party to many of the discussions and central to the security arrangements made between the PA and Israel.

Egyptian duplicity

Throughout the documents, Suleiman in particular is singled out as the point person whom Israeli and American officials could count on to execute their agenda of dividing the Palestinian factions or pressing the PA for greater concessions.

Barely a few months after the January 2006 Palestinian elections that resulted in a Hamas victory, PA leaders were already appealing for assistance in fending off their political opponents. At a meeting between leading Palestinian negotiator Saeb Erekat and US General Keith Dayton, the latter assured the Palestinians that the American administration is committed to reinforcing the PA's Presidential Guard to maintain Mahmoud Abbas' authority in the face of the newly elected Hamas government.

In support of his pledge, Dayton referred to discussions with Suleiman, who committed Egypt, along with Jordan, to providing training and equipment, "even at their own expense".

Later in the year, as the Palestinian factions were engaged in negotiations over the formation of a unity government, a European diplomat told Erekat that the American position on unifying the Palestinians was "prematurely negative". Erekat agreed, adding that Suleiman had also been discouraging of those efforts, saying that they would not work.

In early 2007, as the siege on Gaza had crippling consequences on the lives of Palestinians, negotiators complained that Egyptian leaders were duplicitous, speaking publicly in support of allowing goods into Gaza, but in reality, "it remains blocked on the ground This is a general problem with the Egyptians".

An internal report from April 2007 confirms these suspicions. The Agreement on Movement and Access states: "Although there has been political agreement by Omar Suleiman and President Mubarak on allowing exports through, this agreement has never been translated into operational reality."

Conditions in Gaza only worsened in the months ahead, thanks in large part to the stranglehold imposed by Israel and Egypt. As Hamas assumed sole control of Gaza by preventing a coup attempt by US-backed PA forces, Egypt determined to seal off the border.

In a February 2008 meeting between Ahmed Qurei, a high-ranking PA official, and Tzipi Livni, the then Israeli foreign minister, Qurei relayed the Egyptian position conveyed to him by their leader. "President Mubarak said they'll close down the borders after Sunday and whoever is caught on Egyptian territories will be considered illegal."

The situation came to a stalemate in the months leading up to Israel's December 2008 assault on Gaza that resulted in the deaths of 1,500 Palestinians, most of them civilians. As tensions were heightened, Erekat lamented to his Israeli counterpart that Suleiman was forced to cancel a meeting in the occupied territories. Amos Gilad, the director of Israeli military intelligence, speculated: "Regarding Omar Suleiman, maybe he delayed because he is afraid we will attack while he is here. It will hurt him. He will look like a collaborator."

A tool to pressure Palestinians

The image of Egyptian officials as tools to pressure the Palestinians also emerges out of conversations between US and Palestinian officials. In late 2009, George Mitchell, Barack

Obama's envoy to the region, told Erekat that he had spoken with Suleiman and the two agreed that the PA could unilaterally declare new elections without any input from Hamas.

Furthermore, Mitchell and Suleiman agreed that any agreement would have to permanently eradicate any Hamas presence in the West Bank, while at the same time allowing the PA to resume control of Gaza, terms Hamas was sure to reject. But as Egypt was preparing a document on how the PA should proceed, Erekat assured Mitchell that: "Abu Mazen [Mahmoud Abbas] won't say no to whatever the Egyptians present to him".

Even when it appeared that the Egyptians were attempting to display some degree of autonomy, it became more evident in the documents that external pressure was never too far behind. Only a few weeks later, Erekat complained to US negotiators that Egypt's latest efforts to reconcile the Palestinian factions were straying from the official line. Daniel Rubenstein, the US consul general and chief of mission in Jerusalem, responded: "I can tell you, we did put pressure on the Egyptians. I read the document. It was a disaster."

As Erekat continued to grumble about the PA's weakened position and Egypt's lack of cooperation, General James Jones, the US special envoy for Middle East security, abruptly ended the meeting with his words: "It's insulting. We'll take care of this."

Jones appeared to have lived up to his promise. Only three months later, in January 2010, US negotiator David Hale assured Erekat that in recent talks with Suleiman: "The Egyptians brought ideas similar to our thinking."

In this instance, the US appeared to put pressure on the PA to accept the latest proposals by giving the impression that the US and its allies in the region were unified in their position. Hale further added of the Egyptians: "They talked with Netanyahu and think he is serious."

'Egypt's number two'

Given the critical role that Suleiman has played in advancing US and Israeli objectives, it was no surprise that Mubarak chose to appoint him as vice-president on January 29, a move

rejected by protesters, but reassuring to Egypt's Western patrons. In the leaked documents, Israeli officials were already referring to Suleiman as "Egypt's Number Two" at a time when most observers believed that Mubarak was grooming his son to be succeed him.

Among Western policymakers, it seems Suleiman remains a popular choice to replace Mubarak, as the candidate uniquely suited to maintaining Egypt's current foreign policy, while also addressing domestic grievances expressed by protesters. That remains a distant prospect, given the unlikelihood that the Egyptian opposition would abandon its call to determine the nation's role in regional affairs. But it also demonstrates that, unlike Tunisia, Egypt is far too critical to US objectives in the Middle East to be left to its own devices.

Whatever the outcome in Egypt, it is clear that the recent revelations will have a dramatic impact on the settlement of the Palestinian question. Already weakened by the scandal of the Palestine Papers, Erezkat may now have to do without the support of an Egyptian regime he termed, "our ally, our backbone".

In his first interview as vice-president, Suleiman decried as "unacceptable" what he called "foreign interference" in Egypt's current turmoil. Coming from a regime whose ability to endure through the decades is owed largely to foreign interference, the irony of those words will not be lost on the Egyptian people.

TITLE@In praise of ... Jimmy Carter

DATE@2008 Apr 18

URL@<http://www.theguardian.com/commentisfree/2008/apr/18/usa>

Like the Kennedy Library in Boston, where Gordon Brown makes the main foreign policy speech of his US visit today, most American presidential libraries are monuments to the past. The Carter Centre, near Atlanta, is totally different. Like its begetter, Jimmy Carter, it is focused on the future. The centre thrums with constant activity. Its slogan, "Waging peace, fighting disease, building hope" sums up the work of the most active ex-president the US has ever seen. Mr Carter, now 83, has spent the last quarter-century on a punishing round of conflict resolution, monitoring foreign elections and running medical and other aid programmes in Africa. In 2002 his work won the Nobel peace prize. Six years on there is no letup. Last week Mr. Carter was in Nepal for the elections. This week he is in the Middle East on a peace mission. He met a Hamas delegation from Gaza in Cairo yesterday. Today he moves on to Syria for more talks. These meetings, amid so much recent bloodshed in Gaza, have raised the wrath of the Israeli political establishment against Mr. Carter. And not just the Israeli establishment. Back in the US there have been calls for funding to the Carter Centre to be cut off, while Barack Obama has been forced to say he will not meet Hamas if he becomes president. But Mr. Carter is undeterred - and rightly so. He says Middle East peace will eventually require talks with Syria and talks with Hamas. That is no more than the truth. Today would be a good day for Mr. Brown to say the same thing.

TITLE@Year of hope

DATE@2008 Jan 02

#URL@<http://ekantipur.com/the-kathmandu-post/2008/01/02/editorial/year-of-hope/132912.html>

JAN 01 - The year that has just bade us farewell was one of post-conflict chaos and confusion. Both the two dates announced for the constituent assembly elections came and went without the vote taking place. The past year will be remembered for violent movements, especially in the tarai, and a few but serious conflicts in the western tarai. Crime and political murders spread fast and hard in the tarai. The Young Communist League (YCL) formed last year also set new records in violence and extortion. The year also saw the height of anarchy, impunity and lawlessness. The streets of Kathmandu gained infamy for the most anarchic traffic in the whole world. This was the year Nepal declared itself a Federal Democratic Republic, but the decision needs to be endorsed which will probably happen in 2008. In a nutshell, the year 2007 was a year filled with turmoil.

Living in a country like Nepal, where nothing is certain, it is very hard to be optimistic. However, if we compare the present situation with the period before 2006, we have definitely come a long way. We have fewer people getting killed every day, and the most widely anticipated sustained peace looks achievable. 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. Interestingly, during the final days of 2007, the tarai movement seemed to turn peaceful, the YCL honored its supreme commander's promise and refrained from violent acts, and even the traffic police moved swiftly to restore traffic discipline. It seems we can hope much from the year 2008.

However, the fate of 2008 and beyond depends much on the behavior of our democratic political parties, the Maoists, army, civil society, media, et al. If we all behave responsibly, we will be able to hold the CA elections, accomplish the peace process, restart our economic motor and finally build a new Nepal. But it is definitely not an easy job. The major issue is, of course,

the CA elections; but the problems of the tarai people, janajatis, women, oppressed, marginalized and the like are not so easy to address. More so when we are immersed hopelessly in the politics of corruption, nepotism, favoritism, cronyism and other sorts of ills. If the Maoists do not run away from the elections, if the recently formed and old parties in the tarai live up to their promise 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. We wish you a happy and prosperous New Year.

TITLE@Peace train

DATE@28 DEC 2007 - 03 JAN 2008

#URL@<http://www.nepalitimes.com/issue/2007/12/28/Editorial/14305>

Nepali lefties have always had a flair for pompous rhetoric. Pushpa Kamal Dahal and Babu Ram Bhattarai insist on using a paragraph to say what they can in one sentence.

So we have a 23-point agreement among the seven parties in which the communists commit themselves, once again, to constituent assembly elections.

Nepal has been declared a republic, but it will only take formal effect sometime in the middle of next year after it is ratified by the constituent assembly. But the king is in his palace, still paid a salary by taxpayers money.

The mixed election system has been slightly modified as a face-saving device for the Maoists, but the essence of the earlier arithmetic remains in place. The 23-point agreement is what is already in the interim constitution and in previous accords.

Let's not go into why it took them so long. It's like asking why does the sun set in the west. The important thing is that the peace train which had stopped at a siding (not derailed as some cynics insisted in October when the Maoists left the government) is now on the move again. Essentially, it was never a problem between the leadership of the Maoists and the NC, it was a problem within their parties between the hardliners and moderates. The radicals, of course, are still not satisfied.

Janajatis have also gone along saying the elections are more important for now. Madhesi militant factions have denounced it, but the Forum will take part. The rightwing RJP and RPP dubbed the deal undemocratic. China, India and UNMIN cautiously welcomed the new agreement. The scepticism is understandable because by their past behaviors the seven parties give us no reason to trust them this time.

Polls have been postponed twice. This time, if they fail to have elections by April, the leaders don't just lose face, they lose all legitimacy. Still, from now till April there are three spoilers. Monarchists will try their damndest to not allow elections to happen, an escalation in madesh violence would be a problem, and the wild followers of the Maoists will have to be defanged.

The YCL has harmed the prospects of its own parent party in elections, but the risk is that their excesses will be used by the hard right to wreck elections. Together, the six-party alliance and Maoists must finally prove they are capable of governing.

A government that can't ensure supply of essential commodities, arrange garbage disposal in the capital, guarantee freedom of movement and can't check crime is susceptible to sabotage by society's malcontents.

Nepalis don't want to hear any more speeches. The leaders must see elections as a chance to mend ways and improve their tarnished reputation among the public.

TITLE@CA elections

DATE@2007 Apr 13

#URL@<http://ekantipur.com/the-kathmandu-post/2007/04/13/editorial/ca-elections/106409.html>

APR 12 - The present political hullabaloo is not about whether the CA elections will take place, it's only about when? The Maoists have gone hoarse demanding timely CA polls, 20th June is the formally announced date. UML, as usual, has shown one-upmanship in this regard. UML has forwarded an idea of mobilizing Nepali Army to hold the elections in free and fair manner. Ram Chandra Poudel, Minister for Peace and Reconstruction has asserted that the government is committed to its responsibility of holding elections by mid-June. So, who is against the CA polls? Why have our political parties suddenly felt it so important to reassert their commitment on holding the elections on the said date? Mainly because the election cannot be held on time, and no party wants to take the blame of postponing the election. Reading between the lines, it can be easily conjectured that the CA election is not taking place before October.

The parties have been blaming that the United Nations Mission in Nepal (UNMIN) is attempting to postpone the election just to extend their stay in Nepal. But it is not only the UNMIN, the UK, the USA and other major donors have also shown their doubts about the possibility of holding elections on the given date. Not only the diplomatic circle, even the Election Commission has been emanating a message that EC itself is not prepared to hold the elections. In such a situation, going ahead for the election might not yield the result the people are expecting. Moreover, as the peace and security situation is still weak and the political parties are yet to run campaigns in the villages, the free and fair election is definitely not attainable.

The parties are stressing on the elections mainly because they believe that the delay in holding the elections will provide regressive forces time to foil the elections. Their concern is not unjustified, especially in the backdrop of the history when the postponement of CA polls

after 1950 revolution ultimately resulted in a coup in 1960. However, it is unfortunate that the political parties are trying to label anyone arguing postponement of CA polls as regressive and anti-democratic. It is time we came out candidly about the facts and realities. Just holding the elections for the sake of it will not serve the purpose. We want election that is held in free and fair manner. Obviously, there is a threat from the regressive forces, but if the political parties remain united as in the April movement, no force can change the course of this country. A few months to the interim government to understand the situation and maintain law and order situation before the CA polls will not harm the country. Hence, it's time Koirala, Nepal and Prachanda themselves announced postponement of CA polls citing time constraint. Posted on: 2007-04-11 09:29

TITLE@Is it a Pashtun Question?

AUTHOR@Alistair Burnett

DATE@2008 Sep 12

#

URL@http://www.bbc.co.uk/blogs/theeditors/2008/09/on_the_anniversary_of_the.html

On the anniversary of the September 11th attacks, The World Tonight, had a special edition from Pakistan. Owen Bennett Jones presented the programme from Islamabad while Lyce Doucet reported from Afghanistan.

Seven years on from the attacks in New York and Washington, the key stronghold of groups linked to the Taleban and al-Qaeda is now the wild and remote mountain region straddling the border between Pakistan and Afghanistan. Many call this the new frontline in the battle between western forces, their Afghan and Pakistani allies, and armed Islamic militants.

But there is another way of looking at this region - it is the heartland of the Pashtuns - the tribal people who make up a large element of the population of both Afghanistan and Pakistan, but are resistant to the central authority of both states. The majority of the Taleban are Pashtuns and they have allied themselves to al-Qaeda.

In Afghanistan, American and NATO forces - with Afghan government troops - are involved in an increasingly fierce battle with the Taleban, while in Pakistan 120,000 Pakistani troops are engaged in large scale operations against Taleban fighters and their al-Qaeda allies.

These are the questions we hoped to address in the programme and ones we put to Afghan president Hamid Karzai and Pakistani Foreign Minister, Shah Mahmood Qureshi as well as the British Foreign Secretary, David Miliband.

Do the Pashtuns have specific grievances with the governments in Kabul and Islamabad, which have led to their involvement in the violence? In other words is there a nationalist or tribal element to this conflict as well as a religious one and what does that mean for hopes to end the fighting.

We are not the only ones asking this question - Frederic Grare wrote on this for the Carnegie Endowment for International Peace back in 2006 and Mr. Miliband himself has been blogging on this.

I hope the programme contributed to understanding this complex conflict.

TITLE@The keys to change across the Arab world

AUTHOR@Huguette Labelle

DATE@25 Dec 2011

#URL@<http://www.aljazeera.com/indepth/opinion/2011/12/2011122583047286468.html>

As Tahrir Square fills up again with angry, disillusioned and disappointed citizens, those in power in the Arab world and those who wish to gain power through the ballot box in 2012 should ask themselves one question: How can the next generation of Arab leaders deliver the better world their citizens so courageously demand?

New and old leaders alike must stop the violence and tackle the fundamental problems that created the conditions for inequality, poverty, corruption and repression.

There are three key opportunities for the next generation of parliamentarians to convince their people that they are committed to real reform.

First, transparency, accountability and citizen's rights must take centre stage in the new constitution that they will draft.

Second, rules to keep corruption in check as outlined in the United Nations Convention against Corruption and that have already been ratified by 140 countries, including Tunisia, Egypt and Libya, must be implemented. Enforcing the Convention, which criminalizes bribery and outlines specific laws to curb corruption, would be a strong first step to creating a more equal world. It would also give governments their crucial legitimacy and answer the protests of the people to tackle the corruption that blighted daily life.

And third, all those who abused power in the former regime must be held accountable. Perpetrators should be brought to justice and the assets they stole returned to the country. Clear signals need to be sent that nobody is above the law.

The interim leaders in Egypt can start today by supporting a new constitution that does not put the military above elected politicians and by repealing Law 84 that restricts the role of civil society organizations.

These moves are vitally important. Creating a safe and open space for people and organizations to express their views without government intervention should be a priority across the region. The people were the catalyst for the revolutions and they must be allowed to be the conscience of those who now take over through the ballot box.

That is why removing Law 84 in Egypt is so necessary and why the campaigns against non-governmental organizations and bloggers must stop across the region. Citizens need ways to hold their governments to account.

In Tunisia and Egypt, people queued for hours to cast their ballots, showing their hunger to play a role in shaping their future and a belief that their votes would lead to change. Their representatives have the responsibility not to disappoint them.

The process of drafting new constitution will play an important role in enshrining citizen's rights. But how these constitutional pledges are fulfilled by a country's institutions, its judiciary, laws and its parliament gives the words their power. Indeed, today many of the countries in the Middle East and North Africa have strong legislation. The problem was that it was never implemented.

The United Nations Convention against Corruption, for example, is clear on what is needed to end the impunity that allowed the powerful to abuse their positions to accumulate wealth. It provides a template for comprehensive, well-defined laws to prevent bribery and cronyism. It also provides a legal framework for facilitating the repatriation of stolen assets, a top priority which has become a symbol of the change that has swept the region.

People around the world have a right to hold their governments to account. The ballot box and a vibrant civil society are the best way to do this. But the lessons from recent history are clear: listen to the people, or risk being overtaken by them.

Appendix C Sample of manually annotated editorials and blogs from the corpus

<?xml version="1.0" encoding="utf-8"?>

<Argumentation topic="Maoist's double standard" gist=""
url="http://ekantipur.com/the-kathmandu-post/2007/05/05/editorial/maoists-double-
standard/108572.html" date="2007 May 05" source="The Kathmandu Post">

<Argument id="0">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Negative" persuasion_effect="Yes"
conditional="No" commitment_level="High">A report of the UN Office of the High
Commissioner for Human Rights in Nepal (OHCHR-Nepal), issued last week, manifests the
glaring facts about the CPN-Maoist.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Elaboration"
rhetorical_relation_to="0">In the report the OHCHR-Nepal has starkly said that the Maoist
cadres aren't complying with their party's commitments and are not respecting the rights of
the Internally Displaced Persons (IDPs) to voluntarily and safely return home.</Statement>

<Statement id="2" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Reinforcement"
rhetorical_relation_to="1">The UN rights body has also expressed a serious concern over the
continued existence of the Maoists parallel government and their Kangaroo
courts.</Statement>

<Statement id="3" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes"

conditional="No" commitment_level="High" rhetorical_relation="Reinforcement"
rhetorical_relation_to="2">The scrutiny by the UN rights body conspicuously underlines a
strong message that the Maoists haven't abandoned the culture of intimidation.</Statement>

<Statement id="4" type="Support" support_type="Against" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Elaboration"
rhetorical_relation_to="3">The former rebels non-committal attitude towards their promises
seems to be propelled by their instantaneous fear that the people, who suffered in the hands of
the Maoists, will claim the price of their pain once the rebels lose their control.</Statement>

<Statement id="5" type="Support" support_type="Against" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Reinforcement"
rhetorical_relation_to="3">Equally fair observation is that the Maoists want to continue
political vacuum in the villages.</Statement> <Statement id="6" type="Support"
support_type="Against" exp_type="Opinion" opinion_orientation="Negative"
opinion_strength="Low" persuasion_effect="Yes" conditional="No"
commitment_level="High" rhetorical_relation="Explanation" rhetorical_relation_to="5">For
them, the return of the IDPs means the return of other political parties in the
villages.</Statement>

<Statement id="7" type="Support" support_type="Against" exp_type="Opinion"
opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Contrast"
rhetorical_relation_to="6">But the Maoists don't have the right to obstruct political activities
in villages.</Statement> <Statement id="8" type="Support" support_type="Against"
exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low"
persuasion_effect="Yes" conditional="No" commitment_level="High"
rhetorical_relation="Reinforcement" rhetorical_relation_to="7">Ironically, the Maoist

leadership keeps on reiterating its promise that it will allow the return of the IDPs.</Statement>

<Statement id="9" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Result" rhetorical_relation_to="8">Hence it encapsulates the Maoists carrot and stick approach.</Statement>

</Argument>

<Argument id="1">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="High" conditional="No" commitment_level="High">The return of the IDPs and their property was a cutting issue when the interim government was being formed in March.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Exemplification" rhetorical_relation_to="0">The prime minister and all the political parties had demanded a sincere commitment from the Maoist leadership before including it in the government.</Statement>

<Statement id="2" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration" rhetorical_relation_to="1">The Maoists had agreed in forming a committee of all political parties and immediately facilitating the return of the IDPs and their property, but they didn't abide by their commitment and didn't act accordingly.</Statement>

<Statement id="3" type="Support" support_type="For" exp_type="Fact" persuasion_effect="Yes" conditional="No" commitment_level="Average"

rheterical_relation="Exemplification" rheterical_relation_to="2">The government statistics state that more than 100,000 people in Nepal were displaced during the decade-long insurgency, though it hasn't been independently verified.</Statement>

<Statement id="4" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High" rheterical_relation="Reinforcement" rheterical_relation_to="3">Although the Maoist leadership had announced that it would dissolve its parallel government and Kangaroo courts across the country, they have continued.</Statement>

<Statement id="5" type="Support" support_type="For" exp_type="Fact" persuasion_effect="Yes" conditional="No" commitment_level="High" rheterical_relation="Exemplification" rheterical_relation_to="4">Last week alone, police engaged in a brief scuffle with the Maoist cadres in Kathmandu in an attempt to seize the Maoists vehicle that had the number plate of theso-called Maoist Third Division-4 of PLA.</Statement>

<Statement id="6" type="Support" support_type="For" exp_type="Fact" persuasion_effect="Yes" conditional="No" commitment_level="High" rheterical_relation="Elaboration" rheterical_relation_to="0">Fortunately, officials of the UN Mission in Nepal (UNMIN) took the vehicle and stopped the possible violence.</Statement>

</Argument>

<Argument id="2">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Neutral" opinion_strength="Low" persuasion_effect="Yes" conditional="Yes" commitment_level="High">Since the Maoists have joined the interim government, they also have the responsibility to enforce law and order in the country.</Statement>

<Statement id="1" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Reinforcement" rhetorical_relation_to="0">The Maoist leadership can't play to the gallery.</Statement>

<Statement id="2" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Reinforcement" rhetorical_relation_to="1">It raises a serious question over the Maoists' commitment on democracy and human rights.</Statement>

<Statement id="3" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Reinforcement" rhetorical_relation_to="2">The Post believes that the Maoists double standard is weakening the eight-party alliance.</Statement>

<Statement id="4" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Reinforcement" rhetorical_relation_to="3">The seven political parties in the current coalition government are growing more suspicious of the Maoists just because of the huge gap in their words and practice.</Statement>

<Statement id="5" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Result" rhetorical_relation_to="4">The Post calls on the Maoist leadership to immediately initiate the process of managing a dignified and safe return of the IDPs and respect its commitment to democracy and human rights.</Statement>

</Argument>

</Argumentation>

<Argumentation topic="Egypt and the Palestinian question" gist=""
author="Abdullah Al-Arian"
url="http://www.aljazeera.com/indepth/opinion/2011/02/20112512042978473.html"
date="05 Feb 2011" source="Aljazeera">

<Argument id="0">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="Average">Along with the laundry list of domestic
grievances expressed by Egyptian protesters calling for an end to the regime of Hosni
Mubarak, the popular perception of Egypt's foreign policy has also been a focal point of the
demonstrations.</Statement>

</Argument>

<Argument id="1">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="High">Signs and chants have called on Mubarak to
seek refuge in Tel Aviv, while his hastily appointed vice-president, Omar Suleiman, has been
disparaged as a puppet of the US.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Reinforcement"
rhetorical_relation_to="0">Egypt's widely publicised sale of natural gas to Israel at rock
bottom prices has featured in many refrains emanating from the crowds.</Statement>

</Argument>

<Argument id="2">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High">The widespread view among Egyptians that the regime has served the interests of the West has not been helped by Israel's call for world leaders to support Mubarak, or the apparent unwillingness by American officials to give the protests their full backing.</Statement>

</Argument>

Plummeting status

<Argument id="3">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High">In the shadow of the current cries to topple the Egyptian regime, the Mubarak government has had a tough time keeping its role in international affairs out of public view.</Statement>

</Argument>

<Argument id="4">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High">In the area where Egypt's foreign policy apparatus has served US interests most directly, Israel's security, the Mubarak regime's complicity in the failure to establish a Palestinian state has become widely publicised in recent years.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Elaboration" rhetorical_relation_to="0">Its role in placing the stranglehold on the people of Gaza, in

conjunction with Israel, has seen Egypt's status in the region plummet to a level it has not reached in decades.</Statement>

</Argument>

<Argument id="5">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High">The Palestine Papers, the leaked internal documents of the Palestinian Authority (PA) that were recently exposed by Al Jazeera, provide further confirmation of Egypt's role in the impasse between Israeli and Palestinian negotiators.</Statement>

</Argument>

<Argument id="6">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High">While much of the coverage of the Palestine Papers has focused on the unprecedented concessions offered by Palestinian negotiators, and how swiftly they were spurned by Israeli and American representatives, Egypt's role as an instrument for added pressure stands out from the internal records.</Statement>

</Argument>

<Argument id="7">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High">As the peace process broke down over the past

decade, Egypt was a party to many of the discussions and central to the security arrangements made between the PA and Israel.</Statement>

</Argument>

Egyptian duplicity

<Argument id="8">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High">Throughout the documents, Suleiman in particular is singled out as the point person whom Israeli and American officials could count on to execute their agenda of dividing the Palestinian factions or pressing the PA for greater concessions.</Statement>

</Argument>

<Argument id="9">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High">Barely a few months after the January 2006 Palestinian elections that resulted in a Hamas victory, PA leaders were already appealing for assistance in fending off their political opponents.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Elaboration" rhetorical_relation_to="0">At a meeting between leading Palestinian negotiator Saeb Erekat and US General Keith Dayton, the latter assured the Palestinians that the American administration is committed to reinforcing the PA's Presidential Guard to maintain Mahmoud Abbas' authority in the face of the newly elected Hamas government.</Statement>

</Argument>

<Argument id="10">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High">In support of his pledge, Dayton referred to discussions with Suleiman, who committed Egypt, along with Jordan, to providing training and equipment, "even at their own expense".</Statement>

</Argument>

<Argument id="11">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High">Later in the year, as the Palestinian factions were engaged in negotiations over the formation of a unity government, a European diplomat told Erekat that the American position on unifying the Palestinians was "prematurely negative".</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Elaboration" rhetorical_relation_to="0">Erekat agreed, adding that Suleiman had also been discouraging of those efforts, saying that they would not work.</Statement>

</Argument>

<Argument id="12">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes"

conditional="No" commitment_level="High">In early 2007, as the siege on Gaza had crippling consequences on the lives of Palestinians, negotiators complained that Egyptian leaders were duplicitous, speaking publicly in support of allowing goods into Gaza, but in reality, "it remains blocked on the ground This is a general problem with the Egyptians".</Statement>

</Argument>

<Argument id="13">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High">An internal report from April 2007 confirms these suspicions.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration" rhetorical_relation_to="0">The Agreement on Movement and Access states: "Although there has been political agreement by Omar Suleiman and President Mubarak on allowing exports through, this agreement has never been translated into operational reality."</Statement>

</Argument>

<Argument id="14">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High">Conditions in Gaza only worsened in the months ahead, thanks in large part to the stranglehold imposed by Israel and Egypt.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"

conditional="No" commitment_level="High" rhetorical_relation="Elaboration"
rhetorical_relation_to="0">As Hamas assumed sole control of Gaza by preventing a coup
attempt by US-backed PA forces, Egypt determined to seal off the border.</Statement>

</Argument>

<Argument id="15">

<Statement id="0" type="Opening-statement-thesis" exp_type="Fact"
persuasion_effect="Yes" conditional="No" commitment_level="High">In a February 2008
meeting between Ahmed Qurei, a high-ranking PA official, and Tzipi Livni, the then Israeli
foreign minister, Qurei relayed the Egyptian position conveyed to him by their
leader.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Elaboration"
rhetorical_relation_to="0">"President Mubarak said they'll close down the borders after
Sunday and whoever is caught on Egyptian territories will be considered
illegal."</Statement>

</Argument>

<Argument id="16">

<Statement id="0" type="Opening-statement-thesis" exp_type="Fact"
persuasion_effect="Yes" conditional="No" commitment_level="High">The situation came
to a stalemate in the months leading up to Israel's December 2008 assault on Gaza that
resulted in the deaths of 1,500 Palestinians, most of them civilians.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Elaboration"

rhetorical_relation_to="0">As tensions were heightened, Erekat lamented to his Israeli counterpart that Suleiman was forced to cancel a meeting in the occupied territories.</Statement>

<Statement id="2" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Explanation" rhetorical_relation_to="1">Amos Gilad, the director of Israeli military intelligence, speculated: "Regarding Omar Suleiman, maybe he delayed because he is afraid we will attack while he is here.It will hurt him.He will look like a collaborator."</Statement>

</Argument>

A tool to pressure Palestinians

<Argument id="17">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High">The image of Egyptian officials as tools to pressure the Palestinians also emerges out of conversations between US and Palestinian officials.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration" rhetorical_relation_to="0">In late 2009, George Mitchell, Barack Obama's envoy to the region, told Erekat that he had spoken with Suleiman and the two agreed that the PA could unilaterally declare new elections without any input from Hamas.</Statement>

</Argument>

<Argument id="18">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High">Furthermore, Mitchell and Suleiman agreed that any agreement would have to permanently eradicate any Hamas presence in the West Bank, while at the same time allowing the PA to resume control of Gaza, terms Hamas was sure to reject.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Contrast" rhetorical_relation_to="0">But as Egypt was preparing a document on how the PA should proceed, Erekat assured Mitchell that: "Abu Mazen [Mahmoud Abbas] won't say no to whatever the Egyptians present to him".</Statement>

</Argument>

<Argument id="19">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High">Even when it appeared that the Egyptians were attempting to display some degree of autonomy, it became more evident in the documents that external pressure was never too far behind.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration" rhetorical_relation_to="0">Only a few weeks later, Erekat complained to US negotiators that Egypt's latest efforts to reconcile the Palestinian factions were straying from the official line.</Statement>

<Statement id="2" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration"

rhetorical_relation_to="1">Daniel Rubenstein, the US consul general and chief of mission in Jerusalem, responded: "I can tell you, we did put pressure on the Egyptians. I read the document. It was a disaster."</Statement>

</Argument>

<Argument id="20">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="High">As Erekat continued to grumble about the PA's weakened position and Egypt's lack of cooperation, General James Jones, the US special envoy for Middle East security, abruptly ended the meeting with his words: "It's insulting. We'll take care of this."</Statement>

</Argument>

<Argument id="21">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="Low">Jones appeared to have lived up to his promise.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Neutral" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration" rhetorical_relation_to="0">Only three months later, in January 2010, US negotiator David Hale assured Erekat that in recent talks with Suleiman: "The Egyptians brought ideas similar to our thinking."</Statement>

</Argument>

<Argument id="22">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="Low">In this instance, the US appeared to put pressure on the PA to accept the latest proposals by giving the impression that the US and its allies in the region were unified in their position.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Neutral" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration" rhetorical_relation_to="0">Hale further added of the Egyptians: "They talked with Netanyahu and think he is serious."</Statement>

</Argument>

'Egypt's number two'

<Argument id="23">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High">Given the critical role that Suleiman has played in advancing US and Israeli objectives, it was no surprise that Mubarak chose to appoint him as vice-president on January 29, a move rejected by protesters, but reassuring to Egypt's Western patrons.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration" rhetorical_relation_to="0">In the leaked documents, Israeli officials were already referring to Suleiman as "Egypt's Number Two" at a time when most observers believed that Mubarak was grooming his son to be succeed him.</Statement>

<Statement id="2" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"

conditional="No" commitment_level="High" rhetorical_relation="Elaboration"
rhetorical_relation_to="1">Among Western policymakers, it seems Suleiman remains a popular choice to replace Mubarak, as the candidate uniquely suited to maintaining Egypt's current foreign policy, while also addressing domestic grievances expressed by protesters.</Statement>

<Statement id="3" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Elaboration"
rhetorical_relation_to="2">That remains a distant prospect, given the unlikelihood that the Egyptian opposition would abandon its call to determine the nation's role in regional affairs.</Statement>

<Statement id="4" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Contrast"
rhetorical_relation_to="3">But it also demonstrates that, unlike Tunisia, Egypt is far too critical to US objectives in the Middle East to be left to its own devices.</Statement>

</Argument>

<Argument id="24">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes"
conditional="No" commitment_level="High">Whatever the outcome in Egypt, it is clear that the recent revelations will have a dramatic impact on the settlement of the Palestinian question.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Elaboration"
rhetorical_relation_to="0">Already weakened by the scandal of the Palestine Papers, Erekat

may now have to do without the support of an Egyptian regime he termed, "our ally, our backbone".</Statement>

<Statement id="2" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration" rhetorical_relation_to="1">In his first interview as vice-president, Suleiman decried as "unacceptable" what he called "foreign interference" in Egypt's current turmoil.</Statement>

<Statement id="3" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Result" rhetorical_relation_to="2">Coming from a regime whose ability to endure through the decades is owed largely to foreign interference, the irony of those words will not be lost on the Egyptian people.</Statement>

</Argument>

</Argumentation>

```

<?xml version="1.0" encoding="utf-8"?>

<Argumentation topic="In praise of Jimmy Carter" gist=""
url="http://www.theguardian.com/commentisfree/2008/apr/18/usa" date="2008 April 18"
source="The Guardian">

<Argument id="0">

<Statement id="0" type="Opening-statement-thesis" exp_type="Fact"
persuasion_effect="Yes" conditional="No" commitment_level="Average">Like the Kennedy
Library in Boston, where Gordon Brown makes the main foreign policy speech of his US
visit today, most American presidential libraries are monuments to the past.</Statement>

<Statement id="1" type="Support" support_type="Against" exp_type="Opinion"
opinion_orientation="Neutral" opinion_strength="High" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Contrast"
rhetorical_relation_to="0">The Carter Centre, near Atlanta, is totally different.</Statement>

<Statement id="2" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Exemplification"
rhetorical_relation_to="1">Like its begetter, Jimmy Carter, it is focused on the
future.</Statement>

<Statement id="3" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="Average" rhetorical_relation="Elaboration"
rhetorical_relation_to="2">The centre thrums with constant activity.</Statement>

<Statement id="4" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes"

```


conditional="No" commitment_level="High" rhetorical_relation="Elaboration"
rhetorical_relation_to="3">Its slogan, "Waging peace, fighting disease, building hope" sums
up the work of the most active ex-president the US has ever seen.</Statement>

<Statement id="5" type="Support" support_type="For" exp_type="Fact"
persuasion_effect="Yes" conditional="No" commitment_level="High"
rhetorical_relation="Elaboration" rhetorical_relation_to="4"> Mr. Carter, now 83, has spent
the last quarter-century on a round of conflict resolution, monitoring foreign elections and
running medical and other aid programmes in Africa.</Statement>

<Statement id="6" type="Support" support_type="For" exp_type="Fact"
persuasion_effect="Yes" conditional="No" commitment_level="High"
rhetorical_relation="Cause-Effect" rhetorical_relation_to="5">In 2002 his work won the
Nobel peace prize.</Statement>

<Statement id="7" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes"
conditional="No" commitment_level="High" rhetorical_relation="Reinforcement"
rhetorical_relation_to="6">Six years on there is no letup.</Statement>

<Statement id="8" type="Support" support_type="For" exp_type="Fact"
persuasion_effect="Yes" conditional="No" commitment_level="High"
rhetorical_relation="Exemplification" rhetorical_relation_to="7">Last week Mr Carter was
in Nepal for the elections.</Statement>

<Statement id="9" type="Support" support_type="For" exp_type="Fact"
persuasion_effect="Yes" conditional="No" commitment_level="High"
rhetorical_relation="Exemplification" rhetorical_relation_to="8">This week he is in the
Middle East on a peace mission.</Statement>

<Statement id="10" type="Support" support_type="For" exp_type="Fact"
persuasion_effect="Yes" conditional="No" commitment_level="High"
rhetorical_relation="Exemplification" rhetorical_relation_to="9">He met a Hamas
delegation from Gaza in Cairo yesterday.</Statement>

<Statement id="11" type="Support" support_type="For" exp_type="Fact" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Exemplification" rhetorical_relation_to="10">Today he moves on to Syria for more talks.</Statement>

<Statement id="12" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Explanation" rhetorical_relation_to="11">These meetings, amid so much recent bloodshed in Gaza, have raised the wrath of the Israeli political establishment against Mr Carter.</Statement>

<Statement id="13" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration" rhetorical_relation_to="12">And not just the Israeli establishment.</Statement>

<Statement id="14" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration" rhetorical_relation_to="13">Back in the US there have been calls for funding to the Carter Centre to be cut off, while Barack Obama has been forced to say he will not meet Hamas if he becomes president.</Statement>

<Statement id="15" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Contrast" rhetorical_relation_to="14">But Mr Carter is undeterred - and rightly so.</Statement>

<Statement id="16" type="Support" support_type="Neutral" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Explanation"

rhetorical_relation_to="15">He says Middle East peace will eventually require talks with Syria and talks with Hamas.</Statement>

<Statement id="17" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Reinforcement" rhetorical_relation_to="16">That is no more than the truth.</Statement>

<Statement id="18" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Reinforcement" rhetorical_relation_to="17">Today would be a good day for Mr Brown to say the same thing.</Statement>

</Argument>

</Argumentation>


```

<?xml version="1.0" encoding="utf-8"?>

<Argumentation topic="Year of hope" gist="" author=""
url="http://ekantipur.com/the-kathmandu-post/2008/01/02/editorial/year-of-
hope/132912.html" date="Jan 02, 2008" source="The Kathmandu Post">

<Argument id="0">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="No"
conditional="No" commitment_level="Low">The year that has just bade us farewell was one
of post-conflict chaos and confusion.</Statement>

<Statement id = "1" type="Support" support_type="For" exp_type="Fact"
persuasion_effect="Yes" conditional="No" commitment_level="Low"
rhetorical_relation="Exemplification" rhetorical_relation_to="0">Both the two dates
announced for the constituent assembly elections came and went without the vote taking
place.</Statement>

<Statement id = "2" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="Average" rhetorical_relation="Exemplification"
rhetorical_relation_to="0">The past year will be remembered for violent movements,
especially in the tarai, and a few but serious conflicts in the western tarai.</Statement>

<Statement id = "3" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="Average" rhetorical_relation="Exemplification"
rhetorical_relation_to="0">Crime and political murders spread fast and hard in the
tarai.</Statement>

```

<Statement id = "4" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Exemplification" rhetorical_relation_to="0">The Young Communist League (YCL) formed last year also set new records in violence and extortion.</Statement>

<Statement id="5" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Reinforcement" rhetorical_relation_to="4">The year also saw the height of anarchy, impunity and lawlessness.</Statement>

<Statement id ="6" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Reinforcement" rhetorical_relation_to="5">The streets of Kathmandu gained infamy for the most anarchic traffic in the whole world.</Statement>

<Statement id ="7" type="Support" support_type="For" exp_type="Fact" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Reinforcement" rhetorical_relation_to="6">This was the year Nepal declared itself a Federal Democratic Republic, but the decision needs to be endorsed which will probably happen in 2008.</Statement>

<Statement id = "8" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Result" rhetorical_relation_to="0">In a nutshell, the year 2007 was a year filled with turmoil.</Statement>

</Argument>

<Argument id="1">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="Low">Living in a country like Nepal, where nothing is certain, it is very hard to be optimistic.</Statement>

<Statement id = "1" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Contrast" rhetorical_relation_to="0">However, if we compare the present situation with the period before 2006, we have definitely come a long way.</Statement>

<Statement id = "2" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Elaboration" rhetorical_relation_to="1">We have fewer people getting killed everyday, and the most widely anticipated sustained peace looks achievable.</Statement>

<Statement id = "3" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Reinforcement" rhetorical_relation_to="2">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.</Statement>

<Statement id = "4" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Reinforcement" rhetorical_relation_to="3">Interestingly, during the final days of 2007, the tarai movement seemed to turn peaceful, the YCL honored its supreme commander's promise and refrained from violent acts, and even the traffic police moved swiftly to restore traffic discipline.</Statement>

<Statement id = "5" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="Average" rhetorical_relation="Reinforcement" rhetorical_relation_to="4">It seems we can hope much from the year 2008.</Statement>

</Argument>

<Argument id="2">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Neutral" opinion_strength="Low" persuasion_effect="Yes" conditional="Yes" commitment_level="Low">However, the fate of 2008 and beyond depends much on the behavior of our democratic political parties, the Maoists, army, civil society, media, et al.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" conditional="Yes" commitment_level="Average" rhetorical_relation="Conditional" rhetorical_relation_to="0">If we all behave responsibly, we will be able to hold the CA elections, accomplish the peace process, restart our economic motor and finally build a new Nepal.</Statement>

<Statement id = "2" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Contrast" rhetorical_relation_to="1">But it is definitely not an easy job.</Statement>

<Statement id = "3" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Elaboration" rhetorical_relation_to="2">The major issue is, of course, the CA elections; but the problems of the tarai people, janajatis, women, oppressed, marginalized and the like are not so easy to address.</Statement>

<Statement id = "4" type= "Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" conditional="No" commitment_level="High" rhetorical_relation="Reinforcement" rhetorical_relation_to="3">More so when we are immersed hopelessly in the politics of corruption, nepotism, favoritism, cronyism and other sorts of ills.</Statement>

<Statement id = "5" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" conditional="Yes" commitment_level="High" rhetorical_relation="Conditional" rhetorical_relation_to="0">If the Maoists do not run away from the elections, if the recently formed and old parties in the tarai live up to their promise 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.</Statement>

<Statement id = "6" support="For" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" conditional="No" commitment_level="Low">We wish you a happy and prosperous New Year.</Statement>

</Argument>

</Argumentation>

```

<?xml version="1.0" encoding="utf-8"?>

<Argumentation topic="Peace train" date="28 DEC 2007 - 03 JAN 2008"
url="http://nepalitimes.com/issue/2007/12/28/Editorial/14305" source="Nepali Times">

  <Argument id = "0">

    <Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes"
commitment_level="Low" conditional = "No">Nepali lefties have always had a flair for
pompous rhetoric.</Statement>

    <Statement id = "1" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes"
commitment_level="Average" conditional="No" rhetorical_relation="Exemplification"
rhetorical_relation_to="0">Pushpa Kamal Dahal and Babu Ram Bhattarai insist on using a
paragraph to say what they can in one sentence. </Statement>

  </Argument>

  <Argument id = "1">

    <Statement id = "0" type="Opening-statement-thesis" exp_type="Fact"
persuasion_effect="Yes" commitment_level="High" conditional="No">So we have a 23-
point agreement among the seven parties in which the communists commit themselves, once
again, to constituent assembly elections.</Statement>

  </Argument>

  <Argument id = "2">

    <Statement id = "0" type="Opening-statement-thesis" exp_type="Fact"
persuasion_effect="Yes" commitment_level="High" conditional="Yes"> Nepal has been

```

declared a republic, but it will only take formal effect sometime in the middle of next year after it is ratified by the constituent assembly.</Statement>

<Statement id = "1" type="Support" support_type="Against" exp_type="Fact" persuasion_effect="Yes" commitment_level="Low" conditional="No" rhetorical_relation="Contrast" rhetorical_relation_to="0"> But the king is in his palace, still paid a salary by taxpayers money.</Statement>

</Argument>

<Argument id = "3">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Low" conditional="No">The mixed election system has been slightly modified as a face-saving device for the Maoists, but the essence of the earlier arithmetic remains in place.</Statement>

<Statement id = "1" type="Support" support_type="For" exp_type="Fact" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Elaboration" rhetorical_relation_to="0">The 23-point agreement is what is already in the interim constitution and in previous accords.</Statement>

<Statement id = "2" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Low" conditional="No" rhetorical_relation="Explanation" rhetorical_relation_to="0"> Let's not go into why it took them so long.</Statement>

<Statement id = "3" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Low" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="2"> It's like asking why does the sun set in the west.</Statement>

<Statement id = "4" type="Support" support_type="For" exp_type="Opinion" opinion_orientation = "Positive" opinion_strength = "Low" persuasion_effect="Yes" commitment_level="Low" conditional="No" rhetorical_relation="Elaboration" rhetorical_relation_to="0"> The important thing is that the peace train which had stopped at a siding (not derailed as some cynics insisted in October when the Maoists left the government) is now on the move again.</Statement>

<Statement id = "5" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Low" conditional="No" rhetorical_relation="Elaboration" rhetorical_relation_to="4">Essentially, it was never a problem between the leadership of the Maoists and the NC, it was a problem within their parties between the hardliners and moderates.</Statement>

<Statement id = "6" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Contrast" rhetorical_relation_to="5">The radicals, of course, are still not satisfied.</Statement>

</Argument>

<Argument id = "4">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" commitment_level="Average" conditional="No">Janajatis have also gone along saying the elections are more important for now.</Statement>

<Statement id = "1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="0"> Madhesi militant factions have denounced it, but the Forum will take part.</Statement>

<Statement id = "2" type="Support" support_type= "Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Average" conditional="Low" rhetorical_relation="Contrast" rhetorical_relation_to="0">The rightwing RJP and RPP dubbed the deal undemocratic.</Statement>

<Statement id = "3" type="Support" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Average" conditional="No" rhetorical_relation= "Reinforcement" rhetorical_relation_to="0"> China, India and UNMIN cautiously welcomed the new agreement.</Statement>

<Statement id = "4" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Average" conditional="No" rhetorical_relation="Explanation" rhetorical_relation_to="2"> The scepticism is understandable because by their past behaviours the seven parties give us no reason to trust them this time.</Statement>

</Argument>

<Argument id = "5">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Fact" persuasion_effect="Yes" commitment_level="High" conditional="No">Polls have been postponed twice.</Statement>

<Statement id = "1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" commitment_level="High" conditional="Yes" rhetorical_relation="Conditional" rhetorical_relation_to="0">This time, if they fail to have elections by April, the leaders don't just lose face, they lose all legitimacy.</Statement>

<Statement id = "2" type = "Support" support_type= "For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength = "Average" persuasion_effect="Yes"

commitment_level="High" conditional="No" rhetorical_relation="Explanation"
rhetorical_relation_to="1">Still, from now till April there are three spoilers.</Statement>

<Statement id = "3" type="Support" support_type= "For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes"
commitment_level="High" conditional="No" rhetorical_relation="Exemplification"
rhetorical_relation_to="2">Monarchists will try their damndest to not allow elections to
happen, an escalation in madesh violence would be a problem, and the wild followers of the
Maoists will have to be defanged.</Statement>

<Statement id = "4" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength = "Average" persuasion_effect="Yes"
commitment_level="High" conditional="No" rhetorical_relation="Reinforcement"
rhetorical_relation_to="3">The YCL has harmed the prospects of its own parent party in
elections, but the risk is that their excesses will be used by the hard right to wreck
elections.</Statement>

<Statement id = "5" type="Support" support_type="Against" exp_type="Opinion"
opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes"
commitment_level="Average" conditional="No" rhetorical_relation="Contrast"
rhetorical_relation_to="0">Together, the six-party alliance and Maoists must finally prove
they are capable of governing.</Statement>

<Statement id = "6" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes"
commitment_level="High" conditional="No" rhetorical_relation="Contrast"
rhetorical_relation_to="5">A government that can't ensure supply of essential commodities,
arrange garbage disposal in the capital, guarantee freedom of movement and can't check
crime is susceptible to sabotage by society's malcontents.</Statement>

</Argument>

<Argument id = "6">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" commitment_level="High" conditional="No">Nepalis don't want to hear any more speeches.</Statement>

<Statement id = "1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="0">The leaders must see elections as a chance to mend ways and improve their tarnished reputation among the public.</Statement>

</Argument>

</Argumentation>


```

<?xml version='1.0' encoding='UTF-8'?>

<!DOCTYPE annotation.dtd>

<argumentation title="CA elections" date="2007 Apr 13"
url="http://ekantipur.com/the-kathmandu-post/2007/04/13/editorial/ca-
elections/106409.html">

    <argument id="0">

        <statement id="0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="No"
commitment_level="High" conditional="No">APR 12 - The present political hullabaloo is
not about whether the CA elections will take place, it's only about when?</statement>

        <statement id="1" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"
commitment_level="High" conditional="No" rhetorical_relation="Elaboration"
rhetorical_relation_to="0">The Maoists have gone hoarse demanding timely CA polls, 20th
June is the formally announced date.</statement>

        <statement id="2" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes"
commitment_level="High" conditional="No" rhetorical_relation="Elaboration"
rhetorical_relation_to="1">UML, as usual, has shown one-upmanship in this
regard.</statement>

        <statement id="3" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Neutral" opinion_strength="Low" persuasion_effect="Yes"
commitment_level="average" conditional="No" rhetorical_relation="Explanation"
rhetorical_relation_to="2">UML has forwarded an idea of mobilizing Nepali Army to hold
the elections in free and fair manner.</statement>

```

<statement id="4" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Elaboration" rhetorical_relation_to="3">Ram Chandra Poudel, Minister for Peace and Reconstruction has asserted that the government is committed to its responsibility of holding elections by mid-June.</statement>

<statement id="5" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" rhetorical_relation="Result" rhetorical_relation_to="4">So, who is against the CA polls?</statement>

<statement id="6" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Neutral" opinion_strength="High" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Elaboration" rhetorical_relation_to="5">Why our political parties have suddenly felt it so important to reassert their commitment on holding the elections on the said date?</statement>

<statement id="7" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Explanation" rhetorical_relation_to="6">Mainly because the election cannot be held on time, and no party wants to take the blame of postponing the election.</statement>

<statement id="8" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="7">Reading between the lines, it can be easily conjectured that the CA election is not taking place before October.</statement>

</argument>

<argument id="1">

<statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Low" conditional="No">The parties have been blaming that the United Nations Mission in Nepal (UNMIN) is attempting to postpone the election just to extend their stay in Nepal.</statement>

<statement id="1" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Contrast" rhetorical_relation_to="0">But it is not only the UNMIN, the UK, the USA and other major donors have also shown their doubts about the possibility of holding elections on the given date.</statement>

<statement id="2" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" commitment_level="High" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="1">Not only the diplomatic circle, even the Election Commission has been emanating a message that EC itself is not prepared to hold the elections.</statement>

<statement id="3" type="Support" support_type="Neutral" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Low" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="2">In such a situation, going ahead for the election might not yield the result the people are expecting.</statement>

<statement id="4" type="Support" support_type="Neutral" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="High" persuasion_effect="Yes" rhetorical_relation="Reinforcement" rhetorical_relation_to="3">Moreover, as the peace and security situation is still weak and the political parties are yet to run campaigns in the villages, the free and fair election is definitely not attainable.</statement>

</argument>

<argument id="2">

<statement id="0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" commitment_level="Average" conditional="No">The parties are stressing on the elections mainly because they believe that the delay in holding the elections will provide regressive forces time to foil the elections.</statement>

<statement id="1" type="Support" support_type="Positive" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Explanation" rhetorical_relation_to="0">Their concern is not unjustified, especially in the backdrop of the history when the postponement of CA polls after 1950 revolution ultimately resulted in a coup in 1960.</statement>

<statement id="2" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Contrast" rhetorical_relation_to="1">However, it is unfortunate that the political parties are trying to label anyone arguing postponement of CA polls as regressive and anti-democratic.</statement>

<statement id="3" type="Support" support_type="Neutral" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="2">It is time we came out candidly about the facts and realities.</statement>

<statement id="4" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Neutral" opinion_strength="Low" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="3">Just holding the elections for the sake of it will not serve the purpose.</statement>

<statement id="5" type="Support" support_type="For" exp_type="Opinion" opinion_orientation = "Positive" opinion_strength = "Low" persuasion_effect="Yes" commitment_level = "High" conditional = "No" rhetorical_relation = "Elaboration" rhetorical_relation_to="4"> We want election that is held in free and fair manner.</statement>

<statement id="6" type="Support" support_type="Positive" exp_type = "Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" commitment_level = "High" conditional="No" rhetorical_relation="Elaboration" rhetorical_relation_to="0">Obviously, there is a threat from the regressive forces, but if the political parties remain united as in the April movement, no force can change the course of this country.</statement>

<statement id="7" type="Support" exp_type = "Opinion" opinion_orientation="Positive" opinion_strength = "Low" persuasion_effect="Yes" commitment_level = "High" conditional="No" rhetorical_relation="Elaboration" rhetorical_relation_to="6">A few months to the interim government to understand the situation and maintain law and order situation before the CA polls will not harm the country.</statement>

<statement id="8" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation = "Neutral" opinion_strength="Low" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Cause-Effect" rhetorical_relation_to="7">Hence, it's time Koirala, Nepal and Prachanda themselves announced postponement of CA polls citing time constraint.</statement>

</argument>

</argumentation>

<?xml version="1.0" encoding="utf-8"?>

<Argumentation topic="Is it a Pashtun Question?" editorial_gist="" author="Alistair
Burnett"
url="http://www.bbc.co.uk/blogs/theeditors/2008/09/on_the_anniversary_of_the.html"
date="2008-09-12">

<Argument id="0">

<Statement id="0" type="Opening-statement-thesis" exp_type="Fact">On the
anniversary of the September 11th attacks, The World Tonight, had a special edition from
Pakistan.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Fact"
persuasion_effect="No" conditional="No" rhetorical_relation="Elaboration"
rhetorical_relation_to="0">Owen Bennett Jones presented the programme from Islamabad
while Lyce Doucet reported from Afghanistan.</Statement>

</Argument>

<Argument id="1" >

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Neutral" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="Average">Seven years on from the attacks in New
York and Washington, the key stronghold of groups linked to the Taleban and al-Qaeda is
now the wild and remote mountain region straddling the border between Pakistan and
Afghanistan.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"
commitment_level="Average" conditional="No" rhetorical_relation="Elaboration"> Many
call this the new frontline in the battle between western forces, their Afghan and Pakistani
allies, and armed Islamic militants.</Statement>

</Argument>

<Argument id="2">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Neutral" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="Average">

But there is another way of looking at this region - it is the heartland of the Pashtuns - the tribal people who make up a large element of the population of both Afghanistan and Pakistan, but are resistant to the central authority of both states.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="Average" rhetorical_relation="Elaboration"
rhetorical_relation_to="0">

The majority of the Taleban are Pashtuns and they have allied themselves to al-Qaeda.</Statement>

</Argument>

<Argument id="3">

<Statement id="0" type="Opening-statement-thesis" exp_type="Fact"
persuasion_effect="Yes" conditional="No" commitment_level="High">

In Afghanistan, American and Nato forces - with Afghan government troops - are involved in an increasingly fierce battle with the Taleban, while in Pakistan 120,000 Pakistani troops are engaged in large scale operations against Taleban fighters and their al-Qaeda allies.

</Statement>

</Argument>

<Argument id="4">

<Statement id="0" type="Opening-statement-thesis" exp_type="Fact"
persuasion_effect="No" conditional="No" commitment_level="Low">

These are the questions we hoped to address in the programme and ones we put to Afghan president Hamid Karzai and Pakistani Foreign Minister, Shah Mahmood Qureshi as well as the British Foreign Secretary, David Miliband.

</Statement>

</Argument>

<Argument id="5">

<Statement id="0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="Low"> Do the Pashtuns have specific grievances with the governments in Kabul and Islamabad which have led to their involvement in the violence?</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Negative" opinion_strength="Average" persuasion_effect="Yes"
conditional="No" commitment_level="Average" rhetorical_relation="Reinforcement"
rhetorical_relation_to="0">In other words is there a nationalist or tribal element to this

conflict as well as a religious one and what does that mean for hopes to end the fighting.</Statement>

</Argument>

<Argument id ="6">

<Statement id="0" type="Opening-statement-thesis" exp_type="Fact" persuasion_effect="Yes" conditional="No" commitment_level="Average">

We are not the only ones asking this question - Frederic Grare wrote on this for the Carnegie Endowment for International Peace back in 2006 and Mr Miliband himself has been blogging on this.</Statement>

<Statement id="1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Neutral" opinion_strength="Low" persuasion_effect="No" conditional="No" commitment_level="Low" rhetorical_relation="Conclusion">I hope the programme contributed to understanding this complex conflict.</Statement>

</Argument>

</Argumentation>

<?xml version='1.0' encoding='UTF-8'?>

<argumentation title = "The keys to change across the Arab world" gist = "" author
="Huguette Labelle" date = "25 Dec 2011"
url="http://www.aljazeera.com/indepth/opinion/2011/12/2011122583047286468.html"
source="Aljazeera">

<argument id = "0">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes"
commitment_level="Low" conditional="No">As Tahrir Square fills up again with angry,
disillusioned and disappointed citizens, those in power in the Arab world and those who wish
to gain power through the ballot box in 2012 should ask themselves one question: How can
the next generation of Arab leaders deliver the better world their citizens so courageously
demand?</Statement>

</argument>

<argument id ="1">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes"
commitment_level="High" conditional="No">New and old leaders alike must stop the
violence and tackle the fundamental problems that created the conditions for inequality,
poverty, corruption and repression.</Statement>

</argument>

<argument id ="2">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Neutral" opinion_strength="Average" persuasion_effect="Yes"
commitment_level="High" conditional="No">There are three key opportunities for the next

generation of parliamentarians to convince their people that they are committed to real reform.</Statement>

<Statement id = "1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Elaboration" rhetorical_relation_to="0">First, transparency, accountability and citizen's rights must take centre stage in the new constitution that they will draft.</Statement>

<Statement id = "2" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Elaboration" rhetorical_relation_to="0">Second, rules to keep corruption in check as outlined in the United Nations Convention against Corruption and that have already been ratified by 140 countries, including Tunisia, Egypt and Libya, must be implemented.</Statement>

<Statement id = "3" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" commitment_level="Average" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="2">Enforcing the Convention, which criminalizes bribery and outlines specific laws to curb corruption, would be a strong first step to creating a more equal world.</Statement>

<Statement id = "4" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" commitment_level="Average" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="3">It would also give governments their crucial legitimacy and answer the protests of the people to tackle the corruption that blighted daily life.</Statement>

<Statement id = "5" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Elaboration"

rhetoical_relation_to="0">And third, all those who abused power in the former regime must be held accountable.</Statement>

<Statement id = "6" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" commitment_level="Average" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="5" > Perpetrators should be brought to justice and the assets they stole returned to the country.</Statement>

<Statement id = "7" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="6">Clear signals need to be sent that nobody is above the law.</Statement>

</argument>

<argument id = "3">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Average" persuasion_effect="Yes" commitment_level="Low" conditional="No">The interim leaders in Egypt can start today by supporting a new constitution that does not put the military above elected politicians and by repealing Law 84 that restricts the role of civil society organizations.</Statement>

</argument>

<argument id = "4">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" commitment_level="High" conditional="No">These moves are vitally important.</Statement>

<Statement id = "1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Average" conditional="No" rhetorical_relation="Elaboration" rhetorical_relation_to="0">Creating a safe and open space for people and organizations to express their views without government intervention should be a priority across the region.</Statement>

<Statement id = "2" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="1"> The people were the catalyst for the revolutions and they must be allowed to be the conscience of those who now take over through the ballot box.</Statement>

</argument>

<argument id = "5">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" commitment_level="High" conditional="No">That is why removing Law 84 in Egypt is so necessary and why the campaigns against non-governmental organizations and bloggers must stop across the region.</Statement>

<Statement id = "1" type="Support" support_type="Neutral" exp_type="Opinion" opinion_orientation="Neutral" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Low" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="0">Citizens need ways to hold their governments to account.</Statement>

</argument>

<argument id = "6">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Low" conditional="No">In Tunisia and Egypt, people queued for hours to cast their ballots, showing their hunger to play a role in shaping their future and a belief that their votes would lead to change.</Statement>

<Statement id = "1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" commitment_level="Low" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="0">Their representatives have the responsibility not to disappoint them.</Statement>

<Statement id = "2" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Reinforcement" rhetorical_relation_to="1">The process of drafting new constitution will play an important role in enshrining citizen's rights.</Statement>

<Statement id = "3" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Neutral" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Contrast" rhetorical_relation_to="2"> But how these constitutional pledges are fulfilled by a country's institutions, its judiciary, laws and its parliament gives the words their power.</Statement>

<Statement id = "4" type="Support" support_type="Neutral" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Contrast" rhetorical_relation_to="3">Indeed, today many of the countries in the Middle East and North Africa have strong legislation.</Statement>

<Statement id = "5" type="Support" support_type="Neutral" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes"

commitment_level="High" conditional="No" rhetorical_relation="Contrast"
rhetorical_relation_to="4">The problem was that it was never implemented.</Statement>

</argument>

<argument id = "7">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes"
commitment_level="High" conditional="No">The United Nations Convention against
Corruption, for example, is clear on what is needed to end the impunity that allowed the
powerful to abuse their positions to accumulate wealth.</Statement>

<Statement id = "1" type="Support" support_type="For" exp_type="Opinion"
opinion_orientation="Positive" opinion_strength="Low" persuasion_effect="Yes"
commitment_level="High" conditional="No" rhetorical_relation="Elaboration"
rhetorical_relation_to="0">It provides a template for comprehensive, well-defined laws to
prevent bribery and cronyism.</Statement>

<Statement id = "2" type="Support" support_type = "For" exp_type="Opinion"
opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes"
commitment_level="High" conditional="No" rhetorical_relation="Reinforcement"
rhetorical_relation_to="1"> It also provides a legal framework for facilitating the repatriation
of stolen assets, a top priority which has become a symbol of the change that has swept the
region.</Statement>

</argument>

<argument id = "8">

<Statement id = "0" type="Opening-statement-thesis" exp_type="Opinion"
opinion_orientation="Neutral" opinion_strength="Low" persuasion_effect="Yes"
commitment_level="High" conditional="No">People around the world have a right to hold
their governments to account.</Statement>

<Statement id = "1" type="Support" support_type="For" exp_type="Opinion" opinion_orientation="Positive" opinion_strength="High" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Elaboration" rhetorical_relation_to="0"> The ballot box and a vibrant civil society are the best way to do this.</Statement>

<Statement id = "2" type="Support" support_type="Against" exp_type="Opinion" opinion_orientation="Negative" opinion_strength="Low" persuasion_effect="Yes" commitment_level="High" conditional="No" rhetorical_relation="Contrast" rhetorical_relation_to="1">But the lessons from recent history are clear: listen to the people, or risk being overtaken by them.

</Statement>

</argument>

</argumentation>

Appendix D Top 200 most frequent discourse markers from the corpus

Discourse Marker	Frequency				
And	52796	Finally	1099	Opposite	380
But	21247	Following	1046	Besides	377
Also	8923	Meanwhile	1019	In other words	376
After	7031	So far	981	On the other hand	361
Like	6639	Earlier	970	To continue	358
First	6527	In the past	953	Suddenly	351
Last	6118	After all	934	Rarely	351
While	5864	Sometimes	923	Nevertheless	348
Still	5054	Surely	897	s	
Before	4336	Currently	830	To end	345
Yet	4301	As a result	769	Extremely	339
However	4252	Therefore	756	Undoubtedly	319
That is	3935	Directly	717	y	
Next	2728	Exactly	691	To explain	316
Such as	2724	Otherwise	683	Hence	307
Once	2721	For instance	626	Nonetheless	291
Though	2683	Immediately	624	More	290
Often	2570	And yet	623	important	
Again	2538	Quickly	592	Significantly	289
Second	2434	Equally	579	y	
Always	2298	In particular	574	Above all	286
Instead	2144	Eventually	535	Initially	283
Indeed	1962	At the same time	532	in the future	268
Of course	1888	Generally	525	Absolutely	253
Until	1871	In addition	511	In the first place	246
Although	1812	Unlike	500	Specifically	246
Despite	1792	No doubt	482	In addition to	238
As well as	1726	Similarly	471	In short	237
Especially	1676	Unfortunately	459	Simultaneously	233
Recently	1576	ly		As soon as	224
Later	1567	Obviously	430	Naturally	224
In fact	1483	Previously	423	Slowly	223
Third	1437	Usually	421	At last	222
Thus	1330	In the end	416		
Actually	1245	Repeatedly	410		
for example	1142	Moreover	398		
		At the end	389		

Frequently	218
Likewise	218
Altogether	208
In general	198
In this case	197
In contrast	192
Furthermore	186
By contrast	184
To agree	179
To demonstrate	168
Namely	166
Same as	162
More and more	162
Similar to	153
To close in spite of in the meantime	151
Gradually more	150
Importantly	149
To be sure	147
At first	146
Even so	144
Occasionally	143
Subsequently	143
To admit	142
By now	133
In relation to	122
At present	113
After that	110

Formerly	109
Temporarily	95
To oppose	89
Secondly	82
Instantly	81
On the contrary	79
For one thing	77
Frankly	74
To conclude	73
Most of all	73
In truth	73
Most of all	72
In comparison	69
Consequently	68
To highlight	68
Honestly	68
To some extent	66
To repeat	64
Chiefly	62
the most significant	62
at the beginning	61
Starting with	61
Seldom	58
Equally important	58
Not at all	57
To review	55

As an example	53
To finish	51
Again and again	46
First of all	45
Unquestionably	43
To clarify	41
Least of all	39
Over and over	38
Presently	36
To illustrate	36
On the whole	36
To stress	35
Not long ago	34
Suppose that	34
For all that	34
In most cases	33
The next time	31
Matching	31
All the while	30
Later on	30
In opposition to	30
To reconsider	29

Appendix E Sentiment/Polarity Lexicon

Positive terms		
admire	applause	ever-growing
anti-apartheid	brotherhood	ever-heightening
anti-authoritarian	concile	ever-higher
anti-ballistic	conciliate	ever-mindful
anti-ballistic-missile	contribute	ever-optimistic
anti-boycott	contributed	ever-present
anti-bribery	dazzle	ever-relevant
anti-debt	democracy	ever-rising
anti-fascist	democratic	ever-widening
anti-hate	democratized	faster
anti-imperial	dignified	fellowship
anti-imperialism	double-decker	friendship
anti-imperialist	double-digit	full-fledged
anti-poaching	double-lock	full-grown
anti-poverty	educate	full-hearted
anti-racist	educating	full-scale
anti-sexist	education	full-service
anti-smuggling	educational	full-size
anti-spying	enjoyment	glorious
anti-terror	ever-better	glory
anti-terrorist	ever-cherubic	grateful
anti-Vietnam-war	ever-closer	gratitude
anti-war	ever-faithful	hard-earned
anti-waste	ever-greater	hard-fought

hard-won	long-settled	non-racial
hard-working	long-sought	non-religious
high	long-term	non-sexist
higher	manhood	non-subversive
hope	merge	non-threatening
hopeful	much-admired	non-toxic
hoping	much-anticipated	non-traumatic
improving	much-awaited	non-violently
integrate	much-celebrated	ok
integration	much-cited	okay
law	much-discussed	patch
legal	much-heralded	quasi-genocide
legitimate	much-improved	quasi-illegal
life-affirming	much-loved	quasi-martial
life-death	much-mentioned	quasi-nominal
life-extending	much-needed	reform
life-giving	much-refined	relief
life-long	much-respected	rescue
life-or-death	much-strengthened	respect
life-saving	non-adversarial	respectful
life-supporting	non-aligned	respite
long-anticipated	non-controversial	right
long-awaited	non-corrupt	salvage
long-deserved	non-dangerous	salve
long-established	non-gambler	save
long-expected	non-inflammatory	self-administered
long-heralded	non-interventionary	self-advancement
long-lived	non-interventionist	self-aware
long-predicted	non-nuclear	self-awareness
long-recognized	non-partisan	self-belief
long-serving	non-profit	self-build

self-confessed	self-organise	successful
self-confident	self-policing	super-charge
self-conforming	self-questioning	super-Earth
self-conscious	self-reforming	super-envoy
self-contained	self-regulated	super-fast
self-control	self-regulating	super-lobbyist
self-correcting	self-reliance	super-poodle
self-critical	self-reliant	super-power
self-critique	self-repairing	super-rich
self-defensive	self-replicating	super-secret
self-defined	self-reproducing	super-sexy
self-empowerment	self-restraint	super-size
self-enrichment	self-revelatory	super-summit
self-evident	self-standing	super-wealthy
self-explanatory	self-styled	thankful
self-fund	self-sufficient	thrilled
self-governing	self-support	ultra-cautious
self-government	self-sustained	ultra-cheap
self-healing	self-sustaining	ultra-dynamic
self-help	self-taught	ultra-efficient
self-identified	self-worth	ultra-free
self-identifying	sincere	ultra-local
self-improvement	sisterhood	ultra-precise
self-initiated	sober	ultra-ready
self-investment	still-popular	ultra-refined
self-justifying	still-powerful	ultra-rich
self-knowing	Still-united	unite
self-made	still-young	united
self-made-man	strength	up
self-maintaining	strong	veracious
self-managing	success	victory

war-avoiding	well-funded	well-qualified
well-accepted	well-groomed	well-reasoned
well-acquainted	well-grounded	well-regulated
well-advised	well-hedged	well-rehearsed
well-aimed	well-heeled	well-remembered
well-argued	well-honed	well-researched
well-attended	well-justified	well-resourced
well-behaving	well-laid	well-respected
well-capitalized	well-led	well-rewarded
well-chronicled	well-liked	well-served
well-conceived	well-lit	well-situated
well-coordinated	well-meaning	well-sourced
well-crafted	well-meant	well-stocked
well-defined	well-mobilized	well-structured
well-deserved	well-nigh	well-suited
well-designed	well-off	well-tailored
well-developed	well-oiled	well-thumbed
well-disposed	well-orchestrated	well-trained
well-documented	well-ordered	well-traveled
well-dressed	well-organized	well-trodden
well-earned	well-paid	well-understood
well-entrenched	well-paying	well-versed
well-equipped	well-placed	well-worn
well-evidenced	well-planned	well-written
well-financed	well-poised	well-wrought
well-focused	well-polished	winning
well-fortified	well-prepared	won
well-founded	well-proven	
well-functioning	well-publicized	

Negative terms

abnormalities	anti-Catholic	anti-government
abnormality	anti-censorship	anti-Han
abolishment	anti-Chinese	anti-historical
abolition	anti-choice	anti-immigrant
accuse	anti-Christian	anti-Indian
adamant	anti-college	anti-intellectual
aggressive	anti-colonial	anti-Iran
aggressor	anti-Communist	anti-Iranian
amputated	anti-competitive	anti-Islamic
annexation	anti-Conservative	anti-Islamist
annexing	anti-contraceptive	anti-Israel
anomalies	anti-corporate	anti-labor
anti-access	anti-democratic	anti-market
anti-affirmative	anti-elite	anti-migrant
anti-aircraft	anti-elitist	anti-modern
anti-air-travel	anti-environment	anti-mother
anti-American	anti-environmental	anti-Muslim
anti-Arab	anti-establishment	anti-neoliberal
anti-Berlin	anti-EU	anti-Palestinian
anti-birth	anti-European	anti-political
anti-black	anti-evolutionary	anti-progress
anti-Blair	anti-foreign	anti-protest
anti-Brotherhood	anti-free	anti-Putin
anti-Bush	anti-French	anti-regime
anti-business	anti-gay	anti-religious
anti-capitalist	anti-Georgian	anti-revolutionary

anti-Russian	blood-smeared	cyber-stalked
anti-Saddam	blood-soaked	cyber-warfare
anti-Sarkozy	blood-stained	cynic
anti-Somali	brawl	damaged
anti-Soviet	break-down	danger
anti-state	break-up	debate
anti-Syrian	brutal	defiance
anti-system	closure	dehumanized
anti-Taliban	cold-war-style	deliberate
anti-technology	collapse	deliberately
anti-trust	condemnable	demolish
anti-Turkish	confiscate	demolishment
anti-UN	contradicting	denial
anti-welfare	contradiction	destabilization
anti-western	controversy	destabilize
anti-women	corruption-inflicted	destroy
asylum	corruption-laden	destruction
attack	corruption-plagued	disallowed
away	corruption-ridden	disintegrate
bandh	counter-argument	dismiss
bankruptcy	counter-bid	dismissal
barbarous	counter-claim	double-headed
battle	counter-culture	double-standard
battle-hardened	counter-intuitive	double-whammy
battle-scarred	counter-movement	doubt
battle-weary	counter-proposal	down
beat	counter-revolutionary	drug-addicted
beaten	coup	drug-dealing
below-poverty-line	criticism	drug-induced
bloodbath	criticize	drug-laced
blood-drenched	cyber-bully	drug-manufacturing

drug-producing	forgot	hard-headed
drug-pushing	forgotten	hard-heartedness
drug-related	full-bellied	hard-hitting
drug-running	full-blooded	hard-left
drug-taking	full-force	hard-luck
emptive	full-rigged	hard-nosed
emptively	gang-war-torn	hard-pressed
endangerment	good-for-nothing	hard-right
ever-changing	half-baked	hard-to-prove
ever-contested	half-blooded	hard-to-reach
ever-more-dangerous	half-buried	hard-to-translate
ever-more-stringent	half-burnt	haste
ever-tougher	half-cautioned	hate
ever-wilder	half-dead	hatred
ever-worsening	half-decade	hit
exile	half-decent	ignorance
exiled	half-defused	ignore
extra-constitutional	half-empty	ill-armed
extra-curricular	half-finished	ill-chosen
extra-judicial	half-hearted	ill-considered
extra-legal	half-heartedly	ill-educated
extra-political	half-light	ill-equipped
faulted	half-naked	ill-executed
fearsome	half-party	ill-fitting
fight	half-politician	ill-gotten
fighter	half-readable	ill-health
fighting	half-reformed	ill-informed
fire	half-right	ill-judged
force-fed	hard-and-fast	ill-organised
force-feeding	hard-core	ill-placed
forget	hard-falling	ill-planned

ill-prepared	long-gone	malfunctions
ill-served	long-ignored	malignancy
ill-starred	long-isolated	malnourished
ill-suited	long-oppressed	malnourishment
ill-thought	long-overdue	malnutrition
ill-thought-out	long-suffering	malpractice
ill-thought-through	long-suppressed	malpractices
ill-timed	long-tormented	malware
incorrect	long-unbalanced	manhandle
infertile	long-vanished	militant
infertility	low	military-controlled
inflexible	low-caste	military-dominated
insubordinate	low-energy	military-led
insubordination	lower	military-managed
insurgency	lowest	military-owned
insurgent	low-oxygen	military-ruled
jail	low-paid	military-state
jailed	low-pay	military-style
jeopardy	low-paying	misanthrope
life-destroying	low-profile	misapprehension
life-enhancing	low-skilled	misappropriation
life-prolonging	low-status	misbehavior
loggerhead	Mafia-cum-government	miscalculated
long-coveted	Mafia-like	miscalculation
long-dead	maladies	miscalculations
long-defunct	malady	miscarriage
long-delayed	malaria	miscarriages
long-excluded	malcontents	mischief-making
long-forgotten	Male-dominated	miscommunication
long-fragmented	malfunction	misconceived
long-game	malfunctioning	misconduct

misdeed	misperception	much-trumpeted
misdeeds	misperceptions	much-vaunted
misdeemeanor	misplaced	non-access
misdeemeanors	misquote	non-ceasefire
misdiagnosis	misquoted	non-communicable
misdirected	misreads	non-compliance
miserablist	misrecognition	non-conformist
miseries	mis-reported	non-constitutional
misfired	misreporting	non-credible
misfortunes	misrepresentations	non-defensive
misgoverned	misrule	non-democratic
mishandling	mistaking	non-freedom
mishandlings	mistook	non-funded
mishaps	mistreating	non-inclusive
misidentification	mistreatment	non-licensed
misidentifying	mistrust	non-mainstream
misinformation	misunderstanding	non-negotiable
misinforming	misunderstands	non-renewable
misinterpretations	misunderstood.	non-representative
misinterpreted	misused	non-secular
misjudges	misuses	non-self-governing
misjudgments	misusing	non-state
mislabelling	much-ballyhooed	non-substantial
mislead	much-debated	non-sustainable
misleadingly	much-feared	non-voting
misleads	much-hated	now-abandoned
misled	much-licked	now-convicted
mismanaged	much-maligned	now-divided
mismanagement	much-persecuted	now-fragmented
mismanaging	much-reduced	now-impossible
mismatch	much-touted	now-tired

nuclear-armed	over-imprisonment	plot
nuclear-capable	over-indulgence	Poverty-ridden
nuclear-energy	over-inebriated	pre
nuclear-powered	over-interpreting	premature
opposite	over-investment	prematurely
opposition	over-lawyered	pro-coup
oppress	over-mighty	pro-death
oppressed	over-played	pro-imperial
out-of-character	over-praised	pro-imperialist
out-of-control	over-prescriptive	protest
out-of-date	over-promoted	pro-war
out-of-state	over-pumping	pseudo-apology
out-of-the-blue	over-reached	pseudo-democratic
out-of-touch	over-regulated	pseudo-mystical
over-ambitious	over-reliant	pseudo-police
over-arching	over-reporting	pseudo-revolutionary
over-armed	over-represented	quasi-
Over-baked	over-rouged	quasi-civilian
over-budget	over-sexed	quasi-constitutional
over-centralized	over-sized	quasi-democratic
over-centralizing	over-spending	quasi-double
over-coded	over-spinning	quasi-governmental
over-compensating	over-stated	quasi-independent
over-dedicated	over-stretched	quasi-intellectual
over-doing	over-taxed	quasi-legal
over-excited	overthrow	quasi-legislative
over-exposed	overturn	quasi-military
over-grown	over-warm	quasi-nationalized
over-hasty	over-willingness	quasi-official
over-heated	over-zealous	quasi-papal
over-imagined	peril	quasi-parliamentary

quasi-public	seize	self-harming
quasi-socialist	self-abasement	self-hating
quasi-state	self-absorbed	self-igniting
quasi-withdrawal	self-acquittal	self-image
radical	self-appointed	self-important
radicalize	self-arrogated	self-imposed
rebel	self-asserted	self-imprisonment
rebellion	self-censored	self-indulgent
refusal	self-centered	self-inflicted
reject	self-centeredness	self-loathing
rejection	self-congratulatory	self-martyrdom
removal	self-contradictory	self-obsessed
remove	self-dealing	self-pitying
repress	self-deceiving	self-proclaimed
resist	self-declared	self-professed
resistance	self-deluding	self-propelled
resource-constrained	self-delusion	self-published
resource-consuming	self-denying	self-referential
resource-hungry	self-deprecating	self-regard
resource-intensive	self-deprecatory	self-regarding
resource-strapped	self-depreciating	self-righteousness
risk	self-described	self-righting
rubbish-filled	self-destructed	self-rule
rubbish-strewn	self-devouring	self-sacrifice
savage	self-directed	self-satisfied
scandal-led	self-doubt	self-seeking
scandal-obsessed	self-exile	self-selected
scandal-plagued	self-expression	self-selecting
scandal-ridden	self-fulfilling	self-starving
secession	self-glorifying	self-titled
secessionist	self-harm	severe

severely	too-tight	under-appreciated
short-sighted	ultra-establishment	underclass
short-tempered	ultra-expensive	under-delivering
shut	ultra-hardcore	underdeployment
split	ultra-low	under-developed
stagnating	ultra-nationalist	underdevelopment
still-poor	ultra-orthodox	underemployed
still-rampant	ultra-religious	under-employed
still-squirming	ultra-sensitive	underemployment
still-tenuous	unacceptability	underestimated
still-uncertain	unacceptable	under-estimated
suicide-bomber-style	unaccountability	underestimates
suicide-inducing	unachievable	underestimation
suicide-note	unaddressed	under-funded
super-low	unavailability	underinvestment
super-storm	unavailable	under-investment
super-terrorist	unaware	undernourished
super-wonky	unawareness	undernourishment
suspecting	unbearable	under-nutrition
tear-gas	unbending	under-performance
tear-gassed	uncertain	under-populated
tear-gassing	uncertainties	under-powered
tear-jerking	uncertainty	under-pressure
tensions	unclarity	under-recognized
throw	unconstitutionality	under-regulated
tight	undemocratic	under-remunerated
tighten	underachievement	under-reported
too-brutal	under-achievement	under-representation
too-cranky	underachievers	under-represented
too-hot-to-handle	under-achieving	under-researched
too-narrow	underappreciated	underresourced

under-resourced	unheard	war
under-resourcing	unhygienic	war-battered
under-saving	unilateral	war-crazed
underserved	unilaterally	war-fighting
under-served	unloved	war-filled
under-sized	unlucky	war-fleeing
understaffed	unmanageable	war-game
under-studied	unnecessarily	war-induced
under-supported	unnecessary	war-junky
under-taxed	unpleasantness	war-making
under-the-table	unstable	war-mongering
under-trained	unsuccess	warn
under-used	unsuccessfully	war-provoking
uneducated	unsustainability	war-ravaged
unemployment	unsustainable	war-related
unfair	untidy	war-threatening
unfairly	untimely	war-time
unfairness	untouchability	war-torn
unfaithfulness	untouchable	war-weary
unfashionable	unwelcoming	well-armed
unfavorably	unwell	withdraw
unfeasible	unwitting	withdrawal
unfortunate	unwittingly	withdrew
unfriend	unworkability	worrisome
unfulfilled	vandalize	worst
unhappiness	violence	
unhealthily	violent	

Appendix F Top 200 frequent domain terms (Nepali Times & The Kathmandu Post, 2007-2008)

Frequency, Term			
1805 government	275 issue	215 taken	160 business
1386 Nepal	273 UML	212 police	159 help
1276 people	270 need	212 children	159 force
1180 political	267 minister	211 economic	158 want
1150 Maoists	262 security	210 last	156 set
1098 2008	262 made	209 become	156 free
1091 country	261 peace	207 education	156 already
1072 2007	259 through	205 women	155 monarchy
741 parties	259 come	203 those	155 demand
722 Maoist	258 fact	203 both	154 republic
705 party	257 polls	202 society	154 Dahal
634 kathmandu	251 first	200 support	153 local
634 ekantipur	250 past	199 poor	153 ensure
599 more	248 politics	198 down	153 change
595 time	247 way	196 international	152 forces
567 new	245 Koirala	196 go	150 prices
566 such	235 much	194 number	150 months
550 Nepali	234 constitution	191 major	150 given
540 up	233 NC	190 social	150 development
526 one	232 cadres	189 work	148 provide
513 out	230 national	189 market	148 food
470 over	229 media	189 Kathmandu	148 back
455 CA	229 during	187 percent	146 students
447 state	228 private	186 due	146 making
439 two	228 king	184 under	145 end

434 India	226 violence	183 Prachanda	145 different
398 against	226 prime	182 without	145 better
371 now	223 Minister	182 order	145 again
369 elections	221 situation	181 well	144 part
367 election	220 still	180 sector	144 believe
366 even	220 decision	178 serious	143 world
360 public	220 day	176 Madhesi	143 movement
360 leaders	218 very	176 high	143 future
351 process	218 Indian	173 YCL	143 action
347 years	217 before	171 three	142 Now
345 democracy	216 countries	171 schools	142 good
341 take	142 Constituent	171 demands	128 Prime
323 being	141 economy	169 issues	128 policy
320 rights	140 CPN	168 human	127 responsibility
317 power	139 tarai	168 health	127 place
309 long	139 general	167 hope	127 life
305 make	138 Bhutan	167 going	127 foreign
298 groups	136 press	164 members	127 army
295 right	132 move	164 leadership	125 fuel
290 law	132 interim	164 Assembly	124 run
289 year	132 case	163 Nepalis	124 Editorial
289 many	131 lives	163 needs	123 parliament
287 democratic	130 money	163 community	123 days
284 system	129 matter	162 Post	123 agreement
283 between	129 April	161 problem	122 border

Appendix G Extracted bi-grams from the corpus

Bi-grams			
aberrant behavior	free Middle	increasingly open	massive ordnance
abortive bid	free milk	increasingly	massive
abrupt removal	free mind	opposed	organization
absolute	free money	increasingly opt	massive outbreak
disrespect	free movement	increasingly	massive
absolute	free movement	opting	outbreaks
monarchy	free museums	increasingly	massive overdraft
absurd ideas	free music	organized	massive overhaul
abusive words	free nation	increasingly	massive overhead
acceptable	free nations	outdated	massive
behavior	free negroes	increasingly	paramilitary
accountable	free Negroes	outlandish	massive particle
manner	free news	increasingly	massive payloads
acrimonious	free newspaper	outmoded	massive payments
bickering	free newspapers	increasingly	massive payouts
acute	free not	outraged	massive pension
malnourishment	free nursery	increasingly	massive people
al-Qaeda-run city	free one-off	outspoken	massive
ambiguous stance	free only	increasingly	permanent
ambitious deal	free open-source	overgrown	massive planetary
animal sacrifice	free oxygen	increasingly	massive plume
anomalous	free packet	overstepping	massive police
relationship	free papers	increasingly	massive policy
antagonistic	free parenting	pacified	massive pool
approach	free parking	increasingly paid	massive
anti-American	free pass	increasingly	population
sentiments	free pass	painful	

anti-Putin protests	free pension	increasingly	massive
atrocious	free people	paranoid	populations
conditions	free peoples	increasingly	massive poverty
audacious raid	free phone-ins	penetrable	massive power
awkward	free picture	increasingly	massive
questions	free port	perceive	precautionary
bankrupt regime	free postal	increasingly	massive
belligerent	free power	perceives	prevention
rhetoric	free prescriptions	increasingly	massive
Bhutanese	free press	perplexed	privileges
democracy	free-press	increasingly	massive problem
bird flu	advocacy	persuasive	massive profits
birth pang	Free-primary	increasingly	massive project
bizarre behavior	level	pervading	massive
bland-eye policy	free provision	increasingly	propaganda
blatant income	free range	pervasive	massive property
blatant lies	free-range	increasingly	massive
blue-collar jobs	lectures	picking	proportions
bumptious	free-ranging hens	increasingly	massive
transition	free reign	pinched	proposition
call-bypass	free reign	increasingly plain	massive protest
syndicates	free rein	increasingly	massive protests
callous	free rein	plausible	massive purchase
omnipotence	free ride	increasingly	massive
capricious	free riders	poisoned	quantities
usurpation	free roads	increasingly	massive rally
ceasefire	free run	polarized	massive real-
agreement	free samples	increasingly	estate
chaotic traffic	free school	political	massive
charismatic	free schools	increasingly	redistribution
populism	free schools	politicized	massive reduction

chauvinistic	free sex	increasingly poor	massive reform
journalism	free shares	increasingly	massive reforms
CIA-engineered	free sharing	popular	massive refugee
coup	free show	increasingly	massive
civic mayhem	free societies	populated	registration
civilian bloodshed	free society	increasingly	massive relapse
civil war	free society	porous	massive relapse
class-based rifts	free software	increasingly	massive relief
coercive power	free souls	posing	massive
cold blood	free sounds	increasingly	reparations
complete distrust	free speech	possible	massive
complete	free speech	increasingly	repayments
standstill	free speech	powerful	massive
comprehensive	free spending	increasingly	repercussions
peace	free-spending	practiced	massive
conflict-ridden	households	increasingly	repression
societies	free spirit	precarious	massive rescue
consensual sex	free-spirited	increasingly	massive research
contentious issues	youngsters	precise	massive reset
corruption-	free spirits	increasingly	massive
plagued Nepal	free standing	predicting	retaliation
cost-effective	free-standing	increasingly	massive revenues
intervention	appeal	presented	massive reversal
coup d'etat	free state	increasingly	massive revolt
credulous voters	free State-funded	presenting	massive revolts
crisis-hit market	free states	increasingly	massive rewards
cross-boarder	freest elections	presidential	massive rigging
kidnappings	freest media	increasingly	massive road
cultural clash	freest press	prevalent	massive role
dead constitution	free stuff	increasingly	massive sacrifice
dead end	free superstore	privatized	massive scale

deeply affected	free swimming	increasingly	massive scale
deeply alarmed	free syndicates	professionalized	massive scandal
deeply alienated	free Syria	increasingly	massive scrum
deeply ambivalent	free tea	promising	massive security
deeply angry	free telescopes	increasingly	massive sell-off
deeply	free textbooks	provincial	massive
antipathetic	free themselves	increasingly	sensitivity
deeply anti-	free thinkers	provocative	massive
political	free-thinking	increasingly	separation
deeply anti-	journalists	provoke	massive series
Semitic	free-thinking	increasingly	massive
deeply anxiety-	leadership	pyrrhic	settlement
provoking	free-thinking	increasingly	massive shift
deeply associated	search	questionable	massive shock
deeply attached	free thought	increasingly rabid	massive show
deeply aware	free time	increasingly	massive
deeply betrayed	free-to-air	radical	shrinkage
deeply brutalised	broadcaster	increasingly	massive shut
deeply buried	free-to-air	radicalized	massive skew
deeply care	television	increasingly	massive skills
deeply challenge	free tools	rancid	massive slum
deeply	free tote	increasingly rare	massive slump
challenging	free-to-use NHS	increasingly	massive solidarity
deeply committed	free trade	reactionary	massive
deeply concerned	free trade	increasingly	sovereign
deeply conflicted	free-trade	ready	massive spending
deeply confused	agreement	increasingly real	massive spill
deeply	free-trade	increasingly	massive stake
conservative	agreements	receptive	massive stakes
deeply	free-trade	increasingly	massive state
conservative	consensus	recognize	massive states

deeply considered	free-trade pact	increasingly	massive stimulus
deeply	free trader	recognizes	massive stockpile
contentious	free-trade rules	increasingly	massive stop
deeply contested	free-trade zone	recommend	massive storm
deeply corrupt	free-trading	increasingly	massive storm
deeply corrupted	communists	reflected	massive strategic
deeply	free transfer	increasingly	massive street
counterproductive	free	relevant	massive strike
deeply counter-	transferability	increasingly	massive strikes
productive	free transfers	reliant	massive
deeply criminal	free	increasingly	structures
deeply critical	transportation	religiously-	massive student
deeply cut	free travel	chauvinistic	massive subprime
deeply cynical	free treatment	increasingly	massive subsidies
deeply damaging	free trial	reluctant	massive success
deeply defective	free trip	increasingly	massive sum
deeply depleted	free tuition	relying	massive support
deeply destitute	free university	increasingly	massive swings
deeply destructive	free upload	reordered	massive takeaway
deeply difficult	free use	increasingly	massive task
deeply	free-use era	replaced	massive tax
disappointed	free utilities	increasingly	massive
deeply	free vacation	reprehensible	television
disappointing	free verse	increasingly	massive terror
deeply	free Vitamin	repressive	massive threat
discreditable	free vote	increasingly	massive
deeply dishonest	free voters	resembles	totalitarian
deeply	free votes	increasingly	massive trade
disillusioned	free voting	resentful	massive training
deeply disliked	free water	increasingly	massive transfer
deeply distrust	free web	reshaping	massive transfers

deeply disturbed	free weights	increasingly	massive trial
deeply disturbing	free-wheeling	resort	massive tuition
deeply disturbing	bankers	increasingly	massive
deeply divided	free will	resorted	unemployment
deeply divisive	free women	increasingly	massive
deeply dubious	free world	restive	unemployment
deeply dug	free world	increasingly	massive
deeply	free worship	restless	unfairness
embarrassed	free zone	increasingly	massive uprising
deeply	free zone	restrained	massive urban
embarrassing	hard act	increasingly	massive use
deeply embedded	hard acts	restrictive	massive venture
deeply encrypted	hard-and-fast	increasingly rich	massive
deeply engaged	deadline	increasingly	viewership
deeply English	hard-and-fast	ridiculous	massive
deeply engrained	distinctions	increasingly	violations
deeply enmeshed	hard-and-fast	right-wing	massive violence
deeply entrenched	rules	increasingly risk	massive
deeply entrenched	hard-and-fast	increasingly run	volcanoes
deeply etched	target	increasingly saw	massive voter
deeply familiar	hard answers	increasingly	massive wage
deeply felt	hard argument	sectarian	massive war
deeply flawed	hard artistic	increasingly	massive wealth
deeply flawed	hard ask	secular	massive weapons
deeply foolish	hard	increasingly seek	massive windfall
deeply frustrated	authoritarianisms	increasingly seem	massive
deeply frustrating	hard ball	increasingly seem	worldwide
deeply gloomy	hard bargain	increasingly	military ability
deeply grateful	hard bargain	seems	military abroad
deeply	hard bargaining	increasingly seen	military abuses
handicapped	hard barrier	increasingly sees	

deeply held	hard bet	increasingly	military
deeply held	hard bit	segmented	academies
deeply hesitant	hard blow	increasingly	military academy
deeply honored	hard boards	segregated	military accidents
deeply hostile	hard-boiled	increasingly self-	military
deeply ideological	pragmatists	confident	achievements
deeply ill	hard borders	increasingly	military achieving
deeply illiberal	hard call	sensitive	military act
deeply immoral	hard calls	increasingly	military action
deeply implicated	hard cap	serious	military action
deeply important	hard case	increasingly set	military actions
deeply impressed	hard cases	increasingly	military activities
deeply impressive	hard cash	setting	military activity
deeply indebted	hard cash	increasingly	military actor
deeply infiltrated	hard cheese	settled	military actors
deeply inflected	hard choice	increasingly	military actually
deeply ingrained	hard choices	shabby	military adulation
deeply-ingrained	hard choices	increasingly	military
bipartisan	hard commerce	shaky	advantage
deeply injudicious	hard commitment	increasingly	military
deeply injured	hard	shambolic	adventure
deeply inserted	compromises	increasingly	military
deeply	hard consonants	shape	adventurers
institutionalised	hard convincing	increasingly	military
deeply	hard cop	shaped	adventures
intertwined	hard copy	increasingly	military
deeply	hard core	sharing	adventurism
intoxicating	hard-core birthers	increasingly	military adversary
deeply invested	hard-core cadres	shielded	military advice
deeply involved	hard-core	increasingly shift	military adviser
deeply ironic	communist	increasingly short	military advisers

deeply loathed	hardcore	increasingly	military advocate
deeply lyrical	criminals	shredded	military affairs
deeply marooned	hardcore fans	increasingly	military affairs
deeply	hard-core group	significant	military age
mendacious	hardcore group	increasingly	military-aged
deeply	hard-core	similar	male
metaphysical	instrumentals	increasingly	military-age
deeply militarized	hardcore metal	sinking	males
deeply mindful	hard-core	increasingly	military agencies
deeply misleading	neoliberalism	skeptical	military agencies
deeply mistrustful	hardcore political	increasingly	military agenda
deeply mixed	hard-core politics	skeptical	military agendas
deeply modern	hardcore politics	increasingly	military
deeply moved	hardcore	skewed	aggression
deeply moved	pornography	increasingly small	military agreed
deeply moving	hardcore	increasingly	military
deeply murky	revolutionary	sophisticated	agreements
deeply obnoxious	hardcore social	increasingly	military aid
deeply offensive	hardcore socialist	sought	military airbase
deeply paranoid	hardcore	increasingly	military aircraft
deeply partisan	supporters	sound	military airfield
deeply partisan	hard-core voters	increasingly	military airlift
deeply	hard currency	sounding	military airport
patronizing	hard currency	increasingly	military alliance
deeply penetrated	hard data	speaking	military alliances
deeply personal	hard day	increasingly spent	military allies
deeply personal	hard decision	increasingly	military allies
deeply perverse	hard decisions	standing	military allow
deeply pessimistic	hard deliberation	increasingly stark	military ally
deeply polarized	hard disc	increasingly steep	military ambit
deeply polarizing	hard distinctions		

deeply political	hard documentary	increasingly	military
deeply politicized	hard done	strained	ambitions
deeply pragmatic	hard-done-by	increasingly	military
deeply private	governments	strange	ambitions
deeply	hard drilling	increasingly	military analyst
problematic	hard drive	stressed	military analysts
deeply profound	hard drives	increasingly	military anchor
deeply proud	hard drives	strident	military angle
deeply	hard drug	increasingly	military
questionable	hard-earned	stringent	announcement
deeply	assets	increasingly	military antiques
recalcitrant	hard-earned	strong	military anti-
deeply regrettable	democracy	increasingly	terror
deeply relaxed	hard-earned	subprime	military apparatus
deeply relaxing	dollars	increasingly	military apparatus
deeply religious	hard-earned fruits	subsidized	military appears
deeply reluctant	hard-earned gains	increasingly	military
deeply repellent	hard-earned	substantial	applications
deeply resent	goodies	increasingly	military-appointed
deeply resented	hard-earned	successful	government
deeply resentful	holiday	increasingly	military approach
deeply revealing	hard-earned	suggest	military
deeply rooted	money	increasingly	appropriations
deeply-rooted	hard-earned peace	suggested	military architect
violence	hard-earned	increasingly	military area
deeply rotten	qualification	suggests	military arm
deeply saddened	hard-earned rest	increasingly	military
deeply satisfying	hard-earned	supports	arrangements
deeply scared	salaries	increasingly	military art
deeply scarred	hard-earned	surrounded	military aspects
deeply sceptical	savings		military assault

deeply seared	hard-earned	increasingly	military assault
deeply selective	thriving	suspect	military assaults
deeply shaming	hard-earned	increasingly	military assets
deeply shocking	wages	swirls	military
deeply shrouded	hard economic	increasingly	assistance
deeply social	hardened enough	sympathetic	military
deeply sorry	hardened	increasingly	assistance
deeply stirred	positions	tangled	military
deeply stupid	hard enough	increasingly	assistance
deeply suggestive	harder battle	targeting	military assistant
deeply suspicious	harder drugs	increasingly	military
deeply suspicious	harder drugs	targets	association
deeply symbolic	harder-edged	increasingly taxed	military
deeply	pragmatism	increasingly	assumptions
threatening	harder fought	technological	military/astronaut
deeply tied	harder generation	increasingly	ration
deeply tinged	harder hit	tempted	military attack
deeply touched	harder last	increasingly tense	military attack
deeply	harder now	increasingly	military attacks
transformed	harder nowadays	tenuous	military attacks
deeply	harder nut	increasingly thin	military attempts
transgressed	harder-pressed	increasingly	military attention
deeply troubled	colleagues	thinking	military audience
deeply troubling	harder questions	increasingly	military
deeply	harder reforms	thrust	authorization
uncomfortable	harder sport	increasingly tired	military
deeply	harder stance	increasingly	authorities
unconventional	harder still	tolerant	military
deeply uneasy	harder targets	increasingly	authorities
deeply uneven	harder there	tolerated	military autocrats
deeply unfair	harder time		

deeply	harder times	increasingly	military
unfashionable	harder times	tough	autonomy
deeply	harder today	increasingly	military back
unfortunate	harder-to-track	trailing	military-backed
deeply unhappy	illegal	increasingly	government
deeply unpopular	harder working	transferred	military backers
deeply unreliable	harder-working	increasingly tried	military
deeply unsettling	pupils	increasingly	background
deeply upset	harder-working	troubling	military
deeply upsetting	teachers	increasingly	backgrounds
deeply valued	harder yet	turbulent	military backing
deeply vulnerable	hardest call	increasingly	military backyard
deeply wedded	hardest examples	turned	military bagpipers
deeply worried	hardest	increasingly	military balance
deeply worrying	experiences	turning	military barracks
deeply wounded	hardest hearts	increasingly	military base
deeply woven	hardest hit	tyrannical	military base
deeply wrong	hardest hit	increasingly	military bases
democratic	hardest-hitting	unable	military basing
accountability	analyst	increasingly	military began
democratic	hardest nut	unacceptable	military begins
accountability	hardest part	increasingly	military behavior
democratic	hardest place	unaffordable	military behind
accountability	hardest precisely	increasingly	military beyond
democratic	hardest push	unbelievable	military blockade
achievement	hardest question	increasingly	military blow
democratic	hardest reforms	unclear	military blunders
activism	hardest set	increasingly	military
democratic	hardest step	uncomfortable	bombardments
activists	hardest thing	increasingly	military
democratic age	hardest winter	undermined	bombardments

democratic	hardest word	increasingly	military boots
agenda	hard even	understands	military bordellos
democratically	hard evidence	increasingly	military boss
accountable	hard-faced	undertake	military bosses
democratically	businessmen	increasingly	military box
confirmed	hard-faced	uneasy	military boxes
democratically	society	increasingly	military brass
decide	hard fact	unequal	military breaks
democratically	hard facts	increasingly unfit	military
directed	hard facts	increasingly	breakthrough
democratically	hard-falling	unflattering	military briefers
drafted	victim	increasingly	military brig
democratically	hard fight	united	military brutality
elected	hard fighting	increasingly	military budget
democratically	hard fought	unites	military budget
elected	hard-fought five-	increasingly	military budgets
democratically-	week	universal	military buffers
elected head	hard-fought gains	increasingly	military buildup
democratically-	hard-fought	unmanageable	military build-up
elected leader	transition	increasingly	military build-up
democratically-	hard-fought	unpopular	military bulldozer
elected president	women	increasingly	military bulldozer
democratically	hard game	unpredictable	military
held	hard getting	increasingly	bureaucracy
democratically	hard globalization	unreliable	military
legitimate	hard going	increasingly	bureaucracy
democratically	hard graft	unrepresentative	military-
minded	hard-grafting	increasingly	bureaucratic
democratically	majority	unresponsive	complex
ordained	hard grind	increasingly	military Burmese
	hard harvesting	unstable	military cadre

democratically	hard hat	increasingly	military came
organized	hard hats	unsustainable	military camp
democratically	hard-headed	increasingly	military campaign
ratified	argue	untenable	military campaign
democratically	hard-headed book	increasingly	military
sanctioned	hard-headed	unwelcome	campaigns
democratic	document	increasingly	military camps
alternative	hard-headed	unwilling	military can
democratic	internationalism	increasingly	military
alternative	hard-headed	urgent	capabilities
democratic	management	increasingly used	military
ambitions	hard-headed	increasingly used	capabilities
democratic	pragmatism	increasingly using	military
anomaly	hard-headed	increasingly	capabilities
democratic	realism	vague	military
approval	hard-headed	increasingly	capability
democratic	strategy	valuable	military
argument	hard-hearted	increasingly	capacities
democratic army	nation	varied	military capacity
democratic	hard-helmet work	increasingly	military career
arrangements	hard hit	vibrant	military careers
democratic	hard hit	increasingly	military cargo
aspirations	hard-hit airlines	vicious	military carnage
democratic	hard-hit	increasingly	military casualties
assembly	businesses	views	military casualties
democratic	hard-hit	increasingly	military censor
attempts	communities	violent	military Central
democratic	hard-hit firms	increasingly	military chain
authorities	hard-hit regions	visible	military chains
democratic	hard-hitting	increasingly	military
awakening	campaign	visualized	challenges

democratic	hard-hitting	increasingly vital	military character
backing	speech	increasingly vocal	military
democratic ballot	hard-hitting	increasingly	checkpoints
democratic battle	stories	vociferous	military chief
democratic	hard-hitting	increasingly	military chiefs
beacon	words	volatile	military chiefs
democratic	hard ideal	increasingly votes	military circles
behavior	hard information	increasingly	military circles
democratic birth	hard journey	vulnerable	military-civilian
democratic bodies	hard judgment	increasingly	state
democratic body	hard knocks	wanton	military claim
democratic	hard labor	increasingly wary	military
breakdown	hard labour	increasingly	clampdown
democratic capital	hard landing	watching	military clash
democratic	hard landings	increasingly	military clashes
capitalism	hard language	wayward	military class
democratic case	hard left	increasingly weak	military clients
democratic	hard lesson	increasingly	military cluster
centralism	hard lessons	weeded	military clutches
democratic centre	hard life	increasingly well-	military coalition
democratic	hard line	educated	military code
centrist	hardline anti-	increasingly well-	military coercion
democratic	Maoist	rounded	military
challenges	hard-line camp	increasingly wild	collaboration
democratic	hard-line	increasingly	military
chamber	elements	willing	colleagues
democratic	hard-line faction	increasingly	military
chance	hardline foreign	wireless	columnist
democratic	hard-line leaders	increasingly wish	military
change	hard-line lovers	increasingly	command
		wondering	

democratic	hard-line	increasingly	military
channels	politicians	worried	command
democratic	hard-line politics	increasingly	military
character	hard-line position	worries	commander
democratic China	hard-line pro-war	increasingly	military
democratic	hardline religious	worrying	commanders
choices	hard-line royalists	increasingly	military
democratic citizen	hard-line schools	young	commanders
democratic	hard-line stance	little ability	military
citizenry	hard liquor	little about	commemorations
democratic	hard little	little access	military
citizenship	hard lives	little account	commission
democratic civil	hard look	little account	military
democratic	hard-luck stories	little	commission
civility	hardly able	accountability	military
democratic	hardly absolute	little	commissions
commemoration	hardly absurd	acknowledgement	military
democratic	hardly accords	little adventure	commissions
commitment	hardly addressed	little affiliation	military
democratic	hardly advanced	little affinity	commitment
communities	hardly affected	little agreement	military
democratic	hardly afford	little alarmed	committee
competition	hardly aligned	little allegiance	military
democratic	hardly allayed	little alternative	committees
composition	hardly amused	little apart	military
democratic	hardly answer	little apparent	community
consensus	hardly anyone	little appeal	military
democratic	hardly applies	little appetite	companies
consent	hardly apply	little appetite	military company
democratic	hardly approaches	little appreciated	military
consolidation	hardly argue	little appreciation	competitor

democratic	hardly attractive	little argument	military
constitution	hardly audacious	little arithmetic	competitors
democratic	hardly audible	little army	military complex
constitution	hardly avoid	little attempt	military
democratic	hardly bears	little attention	complicity
contest	hardly beautiful	little attention	military
democratic	hardly became	little attraction	component
contests	hardly begins	little audience	military
democratic	hardly begun	little authority	component
context	hardly behaved	little avatars	military
democratic	hardly boast	little awareness	compound
control	hardly bodes	little background	military computer
democratic	hardly bothered	little baffling	military
convictions	hardly bountiful	little beard	confinement
democratic core	hardly brag	little bearing	military conflict
democratic	hardly bring	little before-the-	military conflict
council	hardly brought	fact	military conflicts
democratic	hardly ceased	little benefit	military
counterparts	hardly ceasing	little benefit	confrontation
democratic	hardly change	little benefit	military
countries	hardly changed	little benefits	confrontation
democratic	hardly checked	little better	military
countries	hardly Christ-like	little better	confrontations
democratic	hardly come	little better-off	military
country	hardly come	little bewildered	confrontations
democratic	hardly comforting	little beyond	military
country	hardly	little bigger	connections
democratic course	commensurate	little bit	military conquest
democratic	hardly	little bite	military
credentials	comparable	little bitterness	conquests
	hardly compares	little black	

democratic	hardly complain	little blame	military
credentials	hardly conceal	little blessing	conquests
democratic	hardly conceded	little blue	military
credibility	hardly confirm	little boats	conscription
democratic crisis	hardly conjure	little boxes	military
democratic	hardly considered	little boy	consensus
culture	hardly constitute	little boy	military
democratic	hardly constitutes	little boys	consequences
culture	hardly	little brain	military
democratic	conventional	little brothers	considerations
cultures	hardly conversant	little burst	military
democratic dawn	hardly convincing	little business	constitution
democratic	hardly count	little capacity	military
deadlock	hardly counts	little capital	construction
democratic debate	hardly covered	little car	military contacts
democratic debate	hardly covers	little care	military contest
democratic debut	hardly delivered	little cash	military continues
democratic	hardly demand	little cause	military continues
decision	hardly describes	little chairs	military
democratic defeat	hardly deserve	little chance	contracting
democratic	hardly deserves	little change	military
deficiencies	hardly designed	little changed	contractors
democratic	hardly deviates	little cheer-up	military
deficiency	hardly devoid	little child	contractors
democratic deficit	hardly differ	little children	military contracts
democratic deficit	hardly different	little choice	military control
democratic	hardly difficult	little choice	military control
delegation	hardly diminish	little clarity	military-
democratic	hardly diminishes	little clearer	controlled spy
deliberation	hardly displayed	little cleverer	military-
	hardly do	little closer	controlled system

democratic	hardly earn	little clue	military controls
demands	hardly embraced	little cocoon	military convoy
democratic	hardly	little coherent	military convoys
development	encompasses	little coin	military
democratic	hardly encourage	little cold	cooperation
dialogue	hardly	little comedy	military
democratic	encouraged	little comfort	cooperation
diction	hardly encourages	little comfort	military co-
democratic	hardly	little command	operation
direction	encouraging	little	military cosmetic
democratic	hardly ends	communication	military costs
directions	hardly engaged	little companions	military council
democratic	hardly enjoy	little company	military council
discourse	hardly enjoying	little compared	military
democratic	hardly exists	little comparison	counterparts
discourse	hardly expect	little complaint	military counter-
democratic	hardly expected	little complicated	weight
discussion	hardly fail	little compromise	military coup
democratic	hardly feels	little concern	military coup
disillusion	hardly find	little concern	military coups
democratic	hardly fired	little concerned	military coups
dissent	hardly fits	little concord	military course
democratic	hardly fix	little confidence	military court
dreams	hardly fulfilling	little connection	military court
democratic	hardly fulfills	little consensus	military courts
dreams	hardly generous	little consequence	military covenant
democratic duty	hardly get	little	military
democratic	hardly getting	consideration	crackdown
dysfunction	hardly glorious	little consolation	military creativity
democratic	hardly go	little contact	military crony
dystopia	hardly got	little context	military culture

democratic	hardly grow	little contrition	military cunning
edification	hardly guarantee	little control	military curfews
democratic	hardly has	little convincing	military currently
election	hardly hedge-	little coordination	military custody
democratic	rows	little coral	military customs
election	hardly held	little cost	military cutbacks
democratic	hardly help	little country	military cuts
elections	hardly helped	little cover	military cuts
democratic	hardly helps	little coverage	military deadlock
elections	hardly imaginable	little creatures	military death
democratic	hardly imminent	little credibility	military deaths
elements	hardly immune	little credit	military
democratic	hardly implies	little cribsheet	decapitation
empowerment	hardly improved	little criticism	military deception
democratic end	hardly incidental	little cross-	military decision-
democratic	hardly indifferent	examination	making
endeavour	hardly inflict	little cultural	military
democratic	hardly	little curiosity	declarations
engagement	intimidated	little currency	military decline
democratic	hardly invited	little cynical	military decree
engagement	hardly justified	little damage	military defeat
democratic	hardly justifies	little danger	military defeats
episode	hardly keep	little dare	military defeats
democratic equity	hardly know	little data	military defeats
democratic era	hardly known	little daylight	Russia
democratic era	hardly large	little debate	military defectors
democratic eras	hardly last	little debates	military defenses
democratic	hardly left	little debt	military defenders
essence	hardly less	little decoding	military defenders
democratic ethos	hardly leveling	little dejected	military delusion
	hardly likely	little democracy	

democratic	hardly limited	little desire	military
exercise	hardly looks	little detail	demobilization
democratic	hardly lusting	little details	military
exercise	hardly made	little deterrent	demolishes
democratic	hardly make	little development	military departure
experience	hardly matter	little deviation	military
democratic	hardly matters	little difference	deployment
experiment	hardly meets	little difference	military
democratic	hardly mentioned	little different	deployments
experiment	hardly missed	little different	military despair
democratic	hardly missing	little difficulty	military detail
experiments	hardly monolithic	little digging	military detail
democratic	hardly more	little	military detainees
explosion	hardly move	disappointing	military detainees
democratic fabric	hardly much	little discipline	military detention
democratic	hardly music	little	military detention
facades	hardly necessary	discrimination	military
democratic	hardly need	little discussion	detentions
factions	hardly needs	little disgust	military
democratic family	hardly negated	little dismayed	deterrence
democratic	hardly negligible	little dismaying	military
fashion	hardly new	little dispute	devastation
democratic feet	hardly noticed	little dissembling	military
democratic	hardly offers	little disturbing	developments
fervour	hardly operating	little done	military
democratic force	hardly	little donkey	developments
democratic forces	outstanding	little dose	military devours
democratic forces	hardly particular	little doubt	military dialogue
democratic form	hardly partisan	little earner	military dictator
democratic forms	hardly persuade	little easier	military dictators
	hardly pick	little easy	military dictators

democratic	hardly places	little education	military
foundation	hardly planning	little educational	dictatorship
democratic	hardly play	little effect	military
framework	hardly popular	little effect	dictatorship
democratic	hardly practical	little effort	military
freedom	hardly preferable	little elaboration	dictatorships
democratic	hardly pretentious	little else	military
freedoms	hardly prevalent	little embarrassed	dictatorships
democratic	hardly probable	little emotion	military
friends	hardly	little emotional	differences
democratic	progressive	little empathy	military
fruition	hardly provides	little emphasis	dimension
democratic	hardly put	little enough	military
functioning	hardly quaking	little enthusiasm	dimensions
democratic	hardly qualifies	little essays	military dinosaurs
functions	hardly quelled	little evidence	military disaster
democratic future	hardly question	little evolution	military disaster
democratic	hardly rational	little excitement	military disasters
futures	hardly realized	little executive	military discipline
democratic gains	hardly reassuring	little experience	military discipline
democratic game	hardly received	little explanation	military
democratic goals	hardly recognize	little exploitative	discretion
democratic	hardly reduced	little exposure	military divisions
governance	hardly refer	little extra	military doctor
democratic	hardly reflects	little extra	military doctrine
governance	hardly refuse	little faith	military
democratic	hardly register	little fanfare	documentation
government	hardly relevant	little fear	military
democratic	hardly remain	little feeling	documents
government	hardly renowned	little fiefdom	military
	hardly represent	little film	dominance

democratic	hardly represents	little financial	military
governments	hardly requires	little firms	dominance
democratic	hardly resolves	little fish	military-
groups	hardly respond	little fishes	dominated State
democratic	hardly retreat	little focus	military draft
history	hardly	little follow	military
democratic house	revolutionary	little follow-up	drawdowns
democratic	hardly rigorous	little food	military dreams
idealism	hardly satisfy	little forward	military drifts
democratic ideals	hardly saying	little freedom	military dusk
democratic ideals	hardly screamed	little fruit	military duty
democratic ideas	hardly seem	little fudging	military/economic
democratic	hardly seems	little funding	elite
identities	hardly served	little fur-ball	military edge
democratic	hardly shown	little further	military effect
ideology	hardly small	little further	military
democratic	hardly snowed	little fuss	efficiency
illiteracy	hardly solve	little gadget	military effort
democratic	hardly sounded	little gain	military elements
impulse	hardly sounds	little galleries	military elite
democratic	hardly speak	little garlic	military elite
impulses	hardly spell	little genius	military elites
democratic	hardly spoiled	little genuine	military embargo
inclusion	hardly squares	little girl	military emerging
democratic India	hardly stand	little girls	military
democratic India	hardly standard	little gold	emphasize
democratic	hardly stellar	little good	military
infrastructure	hardly stood	little governments	encirclement
democratic	hardly stylish	little grip	military
inheritance	hardly subversive	little ground	endeavors
	hardly sufficient	little growth	

democratic	hardly suggests	little guarantee	military
initiation	hardly surprising	little guidance	enforcement
democratic	hardly sweeping	little guilty	military
initiative	hardly symbols	little guy	engagement
democratic	hardly teaches	little guys	military
initiatives	hardly tell	little handicrafts	engagements
democratic	hardly trust	little happiness	military
instinct	hardly unaware	little hard	enterprise
democratic	hardly unique	little harder	military-
instincts	hardly unknown	little harm	entertainment
democratic	hardly	little headway	complex
institution	unprecedented	little headway	military entities
democratic	hardly unusual	little heard	military entity
institutionalization	hardly visible	little-heard radio	military
democratic	hardly walk	little	environments
institutions	hardly want	heartbreaking	military equality
democratic	hardly washed	little heat	military
institutions	hardly welcome	little heed	equipment
democratic	hardly wish	little hells	military
insurance	hardly worked	little help	equipment
democratic intent	hard man	little helper	military
democratic	hard men	little helps	equivalent
intentions	hard money	little hesitation	military era
democratic	hard negotiation	little higher	military
intuition	hard negotiations	little homophobes	escalation
democratic Iran	hard	little hope	military
democratic Iraq	neoliberalism	little horse-race	escalation
democratic issues	hard news	little human	military escorts
democratic justice	hard-nosed	little hurts	military
democratic	businessmen	little ice	essentially
landscape	hard-nosed case	little idea	

democratic	hard-nosed	little ill-defined	military
language	character	little immediate	establishment
democratic	hard-nosed	little impact	military
lawmakers	Conservatives	little impact	establishment
democratic law-	hard-nosed	little import	military
making	industrialist	little importance	establishments
democratic leader	hard-nosed insist	little	military estimates
democratic	hard-nosed	improvement	military ethics
leaders	rightwingers	little incentive	military even
democratic	hard-nosed	little incentive	military
leaders	slugger	little incidents	everywhere
democratic	hard now	little indication	military exchange
leadership	hard numbers	little influence	military
democratic leash	hard nut	little information	exchanges
democratic left	hard parameters	little	military exercise
democratic legacy	hard part	infrastructure	military exercises
democratic	hard-partying	little initiative	military exercises
legitimacy	residents	little insults	military
democratic	hard people	little interaction	expansion
legitimacy	hard pill	little	military
democratic	hard place	interbreeding	expeditions
legitimization	hard place	little interest	military
democratic levers	hard place	little interested	expenditures
democratic	hard places	little internal	military
liberalization	hard policies	little investment	experience
democratic life	hard policy	little iron	military expertise
democratic	hard political	little jingo-	military experts
machinery	hard politics	blogger	military
democratic	hard position	little job	exploitation
mainstream	hard positions	little joy	military exports
	hard pounding	little judicious	military faces

democratic	hard pounding	little justice	military facilities
majority-minority	hard power	little justification	military facility
democratic	hard power	little knowledge	military failure
management	hard power	little known	military failure
democratic	hard pressed	little known	military failures
mandate	hard-pressed	little known	military families
democratic	companies	little-known	military family
mandate	hard-pressed	border	military fatalities
democratic	companies	little-known case	military fear
manner	hard-pressed	little-known	military feels
democratic	exporters	concept	military
market	hard-pressed	little-known fact	feudalism
democratic means	food-producers	little-known party	military fiat
democratic means	hard-pressed	little-known post-	military fight
democratic	governments	9/11	military figures
measures	hard-pressed	little-known	military finally
democratic	households	publication	military firearms
mechanisms	hard-pressed	little-known son	military-first
democratic	journalists	little lady	policy
medium	hard-pressed	little-lamented	military-first
democratic	manufacturers	qualification	response
menace	hard-pressed	little late	military focus
democratic	parents	little later	military footprint
method	hard-pressed	little leaden	military footprint
democratic	police	little leadership	military footprints
midwifery	hard-pressed	little leeway	military force
democratic	population	little leg	military force
milestones	hard-pressed	little legions	military forces
democratic	souls	little legitimacy	military forces
mission	hard-pressed	little less	military/foreign
	states	little leverage	policy

democratic	hard-pressed	little lie	military Former
mobilization	times	little lies	military front
democratic model	hard-pressed	little life	military front
democratic	voters	little light	military frontline
models	hard priorities	little likely	military fuel
democratic	hard pushed	little limp	military fully
movement	hard put	little lines	military functions
democratic	hard put	little live	military functions
movement	hard question	little local	military funding
democratic	hard questioning	little logic	military funerals
movements	hard questions	little love	military gains
democratic	hard rain	little loved	military gains
movements	hard realities	little lower	military garrison
democratic	hard reality	little magazines	military garrison
multiplier	hard reducing	little magic	military General
democratic	hard religion	little mainstream	military generals
muscles	hard reminder	little maneuvering	military generals✗
democratic nation	hard replaced	little margin	military ghosts
democratic	hard research	little massacre	military giant
nations	hard right	little matter	military globally
democratic nature	hard right	little meaning	military goals
democratic NGOs	hard-right	little meat	military goals
democratic	positions	little media	military
nomination	hard road	little medical	governance
democratic norm	hard roads	little military	military
democratic norms	hard rock	little minds	government
democratic norms	hard rocks	little mini-	military-
democratic	hard sacrifices	referendum	government
obligations	hard school	little momentum	nexus
democratic one	hard sciences	little money	military
democratic ones	hard sell	little monsters	governments

democratic ones	hard sell	little moral	military grade
democratic	hard-sell tactics	little more	military-grade
openings	hard simplicities	little more	assault
democratic	hard slog	little motive	military-grade
opinion	hard speech	little motorists	communications
democratic	hard stances	little mountains	military-grade
opposition	hard strategic	little movement	levels
democratic	hard struggle	little mystery	military-grade
optimism	hard struggles	little nebulous	weaponry
democratic order	hard stuff	little necessary	military
democratic order	hard tactics	little need	grandstanding
democratic	hard talk	little need	military gravesite
organisation	hard task	little negative	military gravity
democratic	hard task	little neighbor	military groupies
orientation	hard thing	little nervous	military guarantee
democratic	hard thinking	little new	military guard
outcome	hard time	little news	military habits
democratic	hard times	little note	military
outcome	hard-to-come-by	little notice	hagiography
democratic outfit	resources	little notice	military hand
democratic outfit	hard-to-find spare	little noticed	military hands
democratic	hard toil	little-noticed post	military
oversight	hard-to-pinpoint	little notices	hardliners
democratic	jobs	little objection	military hardware
Pakistan	hard-to-prove	little objective	military
democratic	rapes	little obligation	headquarters
paradigm	hard-to-reach	little odd	military
democratic	areas	little oeuvre	heavyweights
paragons	hard-to-reach	little offering	military
democratic	goal	little opposition	hegemony
parliament		little option	military held

democratic	hard-to-translate	little osmosis	military
participants	periods	little other	helicopter
democratic	hard-to-	little overheated	military help
participation	understand land	little panel	military heroism
democratic parties	hard towards	little parable	military
democratic parties	hard transition	little part	hierarchies
democratic	hard traveling	little parties	military history
partnership	hard trick	little party	military hold
democratic party	hard truth	little passage	military holding
democratic party	hard truths	little patience	military hospitals
democratic path	hard trying	little people	military hostilities
democratic path	hard upbringing	little people	military hubris
democratic patriot	hard way	little peoples	military huts
democratic	hard won	little perspective	military ideas
patronage	hard won	little perverse	military ignore
democratic	hard-won	little	military image-
performance	agreements	philosophical	manager
democratic period	hard-won	little pieces	military
democratic place	campaign	little piglets	imbalance
democratic	hard-won	little place	military impact
platform	commitment	little planet	military impetus
democratic	hard-won extra	little planning	military
policeman	hard-won former	little plants	impotence
democratic	hard-won	little plastic	military
policymaking	freedom	little platform	imprisonment
democratic	hard-won gains	little platoons	military inactivity
politicians	hard-won liberties	little plays	military including
democratic	hard-won peace	little pleasure	military
politics	hard-won	little poached	incompetence
democratic	reputation	little pocket	military incursion
politics	hard-won right	little point	

democratic	hard-won rights	little policy	military
polities	hard-won victory	little political	incursions
democratic polity	hard-won wisdom	little pop	military
democratic polity	hard work	little popular	independence
democratic poll	hard-worked plot	little power	military-industrial
democratic pomp	hard workers	little power	complex
democratic post	hard working	little precious	military industry
democratic	hard-working boy	little presence	military
potential	hard-working	little pressure	infrastructure
democratic	citizens	little pretext	military
potentiality	hard-working	little private-	infrastructure
democratic power	colonists	sector	military initiative
democratic	hard-working	little proclivity	military injuries
powers	councilors	little progress	military inquests
democratic	hard-working	little progress	military
practice	middle	little prominence	installations
democratic	hard-working	little promise	military
practices	people	little prospect	institution
democratic	hard-working	little prospect	military
practices	students	little protection	institutions
democratic	hard-working	little protest	military
precedent	teachers	little provocation	instrument
democratic	hard-work put	little public	military
preference	hard world	little-publicized	intellectuals
democratic	hard yards	poll	military
present	hard year	little push	intelligence
democratic	hard years	little queasy	military interests
present	hard yellow	little quirks	military
democratic	hard yesterday	little rather	intervention
president	increasingly	little rationale	military
	accessible	little real	intervention

democratic	increasingly	little reality	military
pressure	acclaimed	little reason	interventionism
democratic	increasingly	little reassurance	military
pretentions	adamant	little recognized	interventions
democratic	increasingly	little recognition	military
principle	addicted	little recompense	interventions
democratic	increasingly	little recourse	military intranet
principles	adopted	little red	military invasion
democratic	increasingly	little reflection	military invasions
principles	advocate	little regard	military-invested
democratic	increasingly	little regard	projects
procedure	affecting	little relation	military
democratic	increasingly	little relevance	investigation
procedures	affluent	little relevant	military
democratic	increasingly	little reliable	investment
procedures	affordable	little relief	military
democratic	increasingly	little reluctant	involvement
process	aggressive	little remains	military
democratic	increasingly agile	little remarked	irrelevance
process	increasingly	little report	military issues
democratic	agreed	little reported	military jail
processes	increasingly	little reporting	military jargon
democratic	aimless	little research	military jet
processes	increasingly	little resemblance	military journals
democratic	alarmed	little resistance	military judge
programme	increasingly	little resources	military judges
democratic	alienated	little respect	military judgment
progress	increasingly	little respite	military
democratic	aligned	little restraint	juggernaut
project	increasingly	little restriction	military junta
	allow	little retail	military junta

democratic	increasingly	little revolution	military juntas
promise	amorphous	little reward	military jury
democratic	increasingly	little rich	military justice
protest	antagonistic	little ride	military
democratic	increasingly anti-	little risk	launching
protests	government	little risk	military law
democratic	increasingly	little risky	military laws
protests	anxious	little road	military lawyer
democratic	increasingly	little role	military leader
purpose	anything	little room	military leaders
democratic puzzle	increasingly	little room	military
democratic	apparent	little rows	leadership
pyramid	increasingly	little sad	military-led
democratic rather	appears	little safer	attempt
democratic	increasingly	little say	military-led
reasoning	applied	little science	campaign
democratic	increasingly	little scope	military-led order
reconstruction	appreciate	little scorecard	military lesson
democratic	increasingly	little scrutiny	military lessons
referendum	archaic	little scrutiny	military lever
democratic	increasingly	little scuffle	military
reform	arguable	little secret	lieutenants
democratic	increasingly	little secret	military life
reform	artificial	little seems	military links
democratic	increasingly	little sense	military lite
reforms	asserting	little sense	military
democratic	increasingly	little shame	lockdown
reforms	assertive	little shell-	military logic
democratic	increasingly	shocked	military logs
regime	astronomic	little ships	military look
		little short	military loyalists

democratic	increasingly	little sigh	military machine
regimes	attractive	little sign	military machine
democratic	increasingly	little sign	military
renaissance	authoritarian	little significance	machinery
democratic	increasingly	little signs	military man
renewal	autocratic	little similarity	military
democratic	increasingly	little slab	management
renewal	aware	little slack	military mandate
democratic	increasingly bad	little slanderer	military
replacement	increasingly bare	little slower	maneuvers
democratic	increasingly	little smiling	military
representation	based	little snap	manoeuvres
democratic	increasingly	little snow	military
representation	battered	little solace	maneuvering
democratic	increasingly	little space	military massing
republic	became	little spare	military match
democratic	increasingly	little spill	military materials
republics	beginning	little spring	military matters
democratic	increasingly	little stake	military means
responsiveness	behaving	little standing	military means
democratic	increasingly	little startling	military measure
revitalization	believe	little statue	military measures
democratic	increasingly	little statutory	military members
revival	belligerent	little step	military memoirs
democratic	increasingly bent	little stomach	military men
revolts	increasingly	little stranger	military men
democratic	bereft	little street	military methods
revolution	increasingly	little stress	military
democratic	bizarre	little stressed	misadventure
revolution	increasingly blase	little substance	military
		little subtlety	misadventures

democratic	increasingly	little success	military mischief
revolutionary	bloody	little summer	military misrule
democratic-	increasingly blunt	little sunlight	military mission
revolutionary	increasingly blur	little support	military missions
track	increasingly	little support	military mistakes
democratic	blurred	little surplus	military
revolutions	increasingly	little surprise	mobilization
democratic	booby-trapped	little surprise	military
rhetoric	increasingly	little surprised	modernization
democratic right	brazen	little sustenance	military
democratic right	increasingly brief	little sway	movements
democratic rights	increasingly	little switches	military moves
democratic rights	brutal	little sympathy	military muscle
democratic rivals	increasingly	little sympathy	military muscles
democratic roots	burdened	little talk	military nation
democratic route	increasingly	little-talked-about	military nature
democratic rule	bustling	proxy	military nature
democratic rule	increasingly	little tax	military necessity
democratic ruler	calling	little temperance	military needs
democratic rules	increasingly	little theoretical	military news
democratic	capable	little thing	military
science	increasingly cast	little things	notwithstanding
democratic	increasingly	little thinking	military objective
Scotland	catastrophic	little thought	military objects
democratic	increasingly	little thought	military
scrutiny	caused	little threat	obligations
democratic self-	increasingly	little time	military observer
determination	celebrated	little time	military observers
democratic self-	increasingly	little-to-no	military
governance	centred	income	occupation
		little town	military offensive

democratic self-government	increasingly challenged	little trespasser	military offensives
democratic sense	increasingly	little trouble	military offensives
democratic sensibilities	challenges	little trust	military officer
democratic set	increasingly	little turnips	military officers
democratic setting	challenging	little twinkle	military onslaught
democratic setup	increasingly	little unclear	military operation
democratic shortcomings	changing	little uncomfortable	military operation
democratic showpiece	increasingly	little understanding	military operation
democratic socialism	chaotic	little	military operations
democratic socialist	increasingly	little understood	military operations
democratic socialists	cheap	little-understood	military opinion
democratic societies	increasingly	wiggle	military oppression
democratic societies	cheerful	little unless	military option
democratic solution	increasingly	little urgency	military options
democratic solutions	chummy	little use	military order
democratic space	increasingly	little value	military organization
democratic spirit	cinematic	little variety	military organisation
	increasingly	little versus	military outcomes
	circulates	little victory	military outfit
	increasingly clear	little vision	military outpost
	increasingly close	little voter	
	increasingly	little warmth	
	coercive	little warning	
	increasingly	little wars	
	coexists	little water	
	increasingly	little weight	
	colour-blind	little western	
	increasingly come		

democratic spirit	increasingly	little white	military outside
democratic	comes	little wild	military
springtime	increasingly	little willingness	overcomes
democratic	comfortable	little windfall	military
stability	increasingly	little wiser	overconfidence
democratic stake	coming	little wonder	military
democratic stance	increasingly	little word	overflights
democratic	common	little work	military pace
standards	increasingly	little world	military packing
democratic state	communal	little world	military parade
democratic state	increasingly	little worried	military parades
democratic states	competitive	massive abuse	military-
democratic step	increasingly	massive abuses	paramilitary
democratic	complain	massive	collusion
structures	increasingly	achievement	military parity
democratic	complex	massive	military part
struggle	increasingly	advertising	military partially
democratic	comprise	massive aid	military path
struggles	increasingly	massive air	military patrols
democratic style	concentrated	massive	military pause
democratic	increasingly	alienation	military pay
support	concerned	massive amount	military payments
democratic surety	increasingly	massive amounts	military peacocks
democratic	confident	massive anti-	military peers
symbol	increasingly	terror	military
democratic	confined	massive arms	pendulum
system	increasingly	massive army	military pensions
democratic	congruent	massive arrival	military people
system	increasingly	massive	military
democratic	conservative	assassination	performances
systems		massive assault	

democratic	increasingly	massive asset	military
systems	constrained	massive asteroid	perpetrators
democratic	increasingly	massive attacks	military personnel
tempest	contemptible	massive audience	military personnel
democratic theory	increasingly	massive aviation	military
democratic threat	content	massive bailout	philosophy
democratic tide	increasingly	massive bail-outs	military plan
democratic title	contentious	massive balance	military plane
democratic	increasingly	massive ball	military planes
tradition	contribute	massive barrage	military planners
democratic	increasingly	massive base	military planning
tradition	controversial	massive benefit	military plans
democratic	increasingly	massive	military player
traditions	convinced	biography	military players
democratic	increasingly co-	massive Black	military point
transformation	opted	massive	military police
democratic	increasingly	bloodbath	military
transformation	corrupt	massive blow	policeman
democratic	increasingly crave	massive blunders	military policies
transition	increasingly	massive body	military policy
democratic	creaky	massive	military/political
transition	increasingly	bombings	leader
democratic	credible	massive bonuses	military-political
transitions	increasingly	massive bonuses	synergy
democratic	criminalized	massive boost	military politics
transitions	increasingly	massive border	military position
democratic	crippling	massive borrower	military positions
tsunami	increasingly	massive boson	military post
democratic turn	critical	massive brutality	military posts
democratic turn	increasingly cross	massive bubbles	military posture
democratic upper		massive budget	military power

democratic uprising	increasingly crowded	massive build-up	military power
democratic uprisings	increasingly dangerous	massive business	military powerbase
democratic value	increasingly dark	massive campaign	military powerhouse
democratic values	increasingly deal	massive capital	military powers
democratic values	increasingly defiant	massive car	military practice
democratic veneer	increasingly defined	massive case	military precision
democratic venues	increasingly defines	massive cash	military preparations
democratic victories	increasingly demonstrable	massive casualties	military presence
democratic vigour	increasingly denied	massive censorship	military presence
democratic voice	increasingly dependent	massive challenges	military president
democratic voice	increasingly design	massive chamber	military pressure
democratic voices	increasingly despairing	massive change	military preying
democratic vote	increasingly desperate	massive changes	military prison
democratic vote	increasingly destitute	massive chariot	military prisons
democratic wave	increasingly detached	massive civilian	military
democratic way	increasingly design	massive colonialism	production
democratic way	increasingly despairing	massive commitment	military
democratic ways	increasingly desperate	massive computer	professionalism
democratic welfare	increasingly destitute	massive consequences	military
democratic will	increasingly detached	massive conspiracy	programme
democratic wing	increasingly destitute	massive commitment	military
democratic world	increasingly destitute	massive computer	programmes
democratic world	increasingly destitute	massive conflicts	military progress
despotic ruler	increasingly destitute	massive consequences	military projects
devastatingly successful	increasingly destitute	massive consequences	military promises
	increasingly destitute	massive consequences	military
	increasingly destitute	massive consequences	prosecutor

die-hard racists	increasingly	massive	military
extremely active	determined	consumption	prosecutors
extremely adverse	increasingly	massive	military
extremely affluent	dictatorial	coordination	protection
extremely agile	increasingly	massive	military
extremely	difficult	corruption	provocations
angered	increasingly	massive	military prowess
extremely anxious	digital	corruption	military
extremely artful	increasingly dim	massive country	psychiatrist
extremely bad	increasingly dire	massive cousin	military
extremely	increasingly	massive coverage	psychologists
beneficial	direct	massive crack	military
extremely	increasingly	massive crimes	punishment
bewildered	directionless	massive criminal	military purposes
extremely bitter	increasingly	massive critters	military push
extremely blunt	disable	massive cruelty	military
extremely bright	increasingly	massive currency	putschists
extremely bullish	disaffected	massive cut	military quagmire
extremely	increasingly	massive cuts	military rabbinate
burdensome	disastrous	massive cyber	military radios
extremely busy	increasingly	war	military ranks
extremely	disconnected	massive damage	military reach
calculating	increasingly	massive damages	military reaction
extremely	discussed	massive deadly	military readiness
cautious	increasingly	massive debt	military
extremely	discussed	massive debts	realignment
cavalier	increasingly	massive debts	military realities
extremely	disenchanted	massive decisions	military reality
challenging	increasingly	massive defeat	military realm
extremely cheap	disenfranchised	massive defection	military
extremely clear		massive defense	recklessness

extremely close	increasingly	massive deficit	military
extremely	disfigure	massive deficits	reconnaissance
complex	increasingly	massive	military
extremely	dishonored	deforestation	reconstruction
complicated	increasingly	massive de-	military record
extremely	dispersed	industrialization	military records
compromising	increasingly	massive delays	military
extremely	display	massive demand	recruitment
concentrated	increasingly	massive	military reform
extremely	distant	demolitions	military regime
concerned	increasingly	massive	military regime
extremely	distressed	demonstration	military regimes
concerning	increasingly	massive	military regimes
extremely	disturbed	demonstrations	military
confident	increasingly	massive deposits	reinforcements
extremely	disturbing	massive	military relations
confusing	increasingly	destruction	military relations
extremely corrupt	divergent	massive	military
extremely cost-	increasingly	devaluation	relationship
effective	diverse	massive	military
extremely critical	increasingly	difference	relationships
extremely	divided	massive	military
damaged	increasingly	difference	reoccupation
extremely	divided	massive disaster	military report
damaging	increasingly	massive	military
extremely	dominant	disillusionment	representatives
dangerous	increasingly	massive	military
extremely	dominate	displacement	repression
dangerous	increasingly	massive	military
extremely delicate	dominated	disruption	reputation
		massive distances	military research

extremely	increasingly	massive distortion	military
demanding	dominated	massive	resolution
extremely devious	increasingly	distribution	military resolve
extremely	doubtful	massive drain	military resource
difficult	increasingly	massive drug	military resources
extremely	drawn	massive	military response
dignified	increasingly	dysfunction	military response
extremely dirty	driven	massive	military
extremely	increasingly dry	earthquake	responsibilities
disabled	increasingly	massive	military
extremely	dubious	earthquakes	retrenchment
disappointing	increasingly	massive effort	military reversals
extremely	dysfunctional	massive efforts	military review
disillusioned	increasingly	massive	military rise
extremely	dystopian	electricity	military rival
disturbing	increasingly	massive elephant	military rivalry
extremely	effective	massive	military rivals
disunited	increasingly elite	embezzlement	military role
extremely	increasingly	massive	military rule
divisive	embracing	emissions	military rule
extremely	increasingly	massive	military rule
doubtful	emerging	endorsements	military ruler
extremely	increasingly	massive energy	military rulers
dubious	empowered	massive	military security
extremely	increasingly	engineering	military sense
effective	entering	massive enough	military
extremely	increasingly	massive ethanol	separation
embarrassing	entrenched	massive	military service
extremely evil	increasingly	evacuations	military service
extremely	erratic	massive event	military services
exciting		massive events	military setback

extremely	increasingly	massive exercise	military
expensive	essential	massive exodus	showdown
extremely fit	increasingly	massive expanses	military side
extremely flexible	establishment	massive	military siege
extremely focused	increasingly	expansion	military
extremely foreign	estranged	massive	significance
extremely	increasingly	expenditure	military signing
fortunate	evident	massive	military site
extremely fragile	increasingly	exploitation	military sites
extremely funny	excluded	massive	military situation
extremely general	increasingly	expression	military
extremely good	executed	massive failure	skirmishes
extremely grateful	increasingly	massive failures	military slaughter
extremely grave	expanded	massive farm	military solidarity
extremely grim	increasingly	massive film	military solution
extremely hard	expected	massive finance	military solution
extremely harsh	increasingly	massive fine	military solutions
extremely heavy	expensive	massive fines	military sought
extremely high	increasingly	massive	military sources
extremely hostile	exposed	firepower	military spaceship
extremely ill-	increasingly	massive floods	military spender
judged	exposing	massive flow	military spending
extremely	increasingly	massive	military spending
immoral	expressing	followings	military spends
extremely	increasingly face	massive force	military spheres
important	increasingly	massive force	military spheres
extremely indirect	failing	massive fraud	military
extremely	increasingly fall	massive funding	spokesman
infectious	increasingly false	massive	military
extremely	increasingly	fundraising	spokesmen
interesting	familiar	massive funds	

extremely knowledgeable	increasingly fatalistic	massive gambles	military
extremely large	increasingly fed	massive gas	spokesperson
extremely late	increasingly feel	massive giveaway	military squads
extremely likely	increasingly felt	massive	military staff
extremely limited	increasingly	government	military stalemate
extremely loath	fierce	massive grids	military stalemate
extremely	increasingly find	massive growth	military stance
localised	increasingly	massive health	military standoff
extremely low	flattened	massive heart	military state
extremely lowball	increasingly flirts	massive helping	military-state
extremely loyal	increasingly	massive history	power
extremely	focusing	massive hoax	military status
lucrative	increasingly	massive hole	military step
extremely	folding	massive	military still
misleading	increasingly	homophobia	military stockade
extremely modest	forthright	massive housing	military strategic
extremely murky	increasingly	massive human	military strategies
extremely	found	massive	military strategy
nauseating	increasingly	ignorance	military strategy
extremely odd	fractious	massive	military strength
extremely opaque	increasingly	imbalance	military strength
extremely	fragile	massive	military strike
partisan	increasingly	imbalances	military strikes
extremely	frame	massive	military strikes
patriarchal	increasingly	immigration	military structure
extremely	frequent	massive impact	military struggle
pernicious	increasingly	massive import	military study
extremely	fretting	massive income	military style
politicised	increasingly	massive	military-style
extremely poor	frustrated	incompatibility	assault
		massive increase	

extremely popular	increasingly	massive industry	military-style
extremely	gaining	massive	compound
positive	increasingly	inequalities	military-style
extremely	geared	massive	tactics
powerful	increasingly	inequities	military-style
extremely	getting	massive infernos	training
problematic	increasingly give	massive influence	military success
extremely	increasingly	massive influx	military success
promising	giving	massive	military successes
extremely	increasingly	information	military
pronounced	global	massive	superiority
extremely prudent	increasingly	injections	military
extremely rare	globalised	massive	superiority
extremely real	increasingly good	instability	military superiors
extremely	increasingly got	massive	military
reasonable	increasingly	investigation	superpower
extremely	grade-conscious	massive	military
reckless	increasingly	investment	superpower
extremely remote	grainy	massive	military
extremely restless	increasingly grim	investments	superpowers
extremely rich	increasingly	massive Israeli	military supplies
extremely	grounded	massive job	military supply
rigorous	increasingly hang	massive	military support
extremely sad	increasingly hard	journalism	military support
extremely salient	increasingly harsh	massive law	military
extremely	increasingly have	massive lay-offs	supporting
sensible	increasingly	massive loan	military
extremely	heard	massive loans	suppression
sensitive	increasingly	massive loss	military
extremely serious	heated	massive losses	supremacy
extremely short		massive luxury	military surge

extremely short-sighted	increasingly heavier	massively armed	military surge
extremely significant	increasingly heavy	massively attacks	military surges
extremely sketchy	increasingly held	massively buying	military surveyors
extremely slippery	increasingly heretical	massively	military survival
extremely small	increasingly high-profile	complex	military system
extremely sophisticated	increasingly holding	massively corrupt	military tactician
extremely sorry	increasingly homogenous	massively	military tactics
extremely strong	increasingly hostile	counterproductive	military tactics
extremely stupid	increasingly hot	massively	military takeover
extremely subversive	increasingly hungry	damaging	military talks
extremely successful	increasingly ice-free	massively	military tapes
extremely suspicious	increasingly identify	destructive	military target
extremely talented	increasingly illiberal	massively	military targets
extremely tepid	increasingly impatient	different	military tasks
extremely tight	increasingly impersonal	massively	military
extremely tragic		disposed	military technicians
extremely troubling		massively	military
extremely turbulent		egotistical	technology
extremely uncomfortable		massively	military terms
		encouraged	military terms
		massively	military test
		expanding	military theatres
		massively	military thing
		extended	military threat
		massively failed	military threat
		massively funded	military threats
		massively	military ties
		hierarchical	military trade
		massively	military-trained
		important	thugs
			military trainers

extremely	increasingly	massively	military training
undesirable	implausible	increase	military trains
extremely	increasingly	massively-	military transit
unethical	important	increased	military transport
extremely	increasingly	spending	military tree
unfortunate	imports	massively	military trial
extremely	increasingly	inexperienced	military trials
unpleasant	inadequate	massively inflated	military trials
extremely	increasingly	massively inhibit	military tribunal
unpopular	incensed	massively	military tribunals
extremely	increasingly	invested	military tribunals
unrealistic	includes	massively	military triumph
extremely	increasingly	irresponsible	military troops
unreliable	incompatible	massively kitsch	military turning
extremely	increasingly	massively	military turns
unspecific	independent	outnumbered	military tweet
extremely unwise	increasingly	massively popular	military tweets
extremely useful	indistinguishable	massively raised	military types
extremely valid	increasingly	massively	military umbrella
extremely	industrialized	regressive	military
variable	increasingly	massively	uncertainty
extremely vigilant	ineffective	sensitized	military
extremely violent	increasingly	massively	undertaking
extremely volatile	ineffectual	strengthen	military uniform
extremely	increasingly	massively swept	military uniform
vulnerable	inefficient	massively tighten	military unit
extremely wary	increasingly inept	massively	military units
extremely weak	increasingly	underestimated	military
extremely	influencing	massively	unpleasantness
weighty	increasingly	underfunded	military use
	influential		military use

extremely well-qualified	increasingly inspire	massively understating	military uses
extremely wide	increasingly	massive machine	military velour
extremely wide-ranging	institutionalized	massive majority	military vehicles
extremely worried	increasingly	massive manhunt	military venture
free access	intense	massive media	military ventures
free	increasingly	massive	military versus
accommodation	interacting	migration	military veterans
free admission	increasingly	massive mistakes	military victory
free adoption	international	massive misuse	military victory
free adults	increasingly	massive	military violence
free advertising	interventionist	mobilisations	military virility
free advice	increasingly	massive move	military virility
free agent	intolerant	massive	military virtually
free agents	increasingly	movements	military voices
free-and-easy	introverted	massive	military
lending	increasingly	nationwide	vulnerabilities
free answers	inventive	massive non-cooperation	military wants
free apps	increasingly	cooperation	military war
free association	involved	massive number	military warlord
free atmosphere	increasingly	massive numbers	military
free bailout	irrational	massive obstacle	whistleblowers
free ball	increasingly	massive offensive	military wing
free banking	irrelevant	massive offshore	military wing
free beer	increasingly	massive oil	military
free black	isolated	massive online	withdrawal
free blacks	increasingly		military withdrew
free board	lacked		military world
free books	increasingly large		military zone
free bosses	increasingly		military zone
	lawless		significant
			number
			significant part

free buildings	increasingly less		stark contrast
free bus	increasingly		strategically
free business	lethal		important
free candy	increasingly		too keen
free capital	likely		too late
free care	increasingly limit		too little
free cash	increasingly little		too many
free cataract	increasingly live		too much
free channels	increasingly		too narrow
free childcare	longer		too old
free chips	increasingly look		too young
free choice	increasingly look		undue influence
free choices	increasingly looks		vehemently
free churches	increasingly		denied
free citizen	losing		well educated
free citizens	increasingly lost		well established
free citizens	increasingly loud		widespread
free coffee	increasingly		support
free communities	lucrative		
free community	increasingly made		
free competition	increasingly		
free condoms	makes		
free content	increasingly		
free contestation	manic		
free contraception	increasingly		
free	manifest		
contraceptives	increasingly		
free cost	marginalised		
free country	increasingly		
free culture	massive		
free data			

free deals	increasingly		
free debate	materialistic		
free delivery	increasingly		
free dessert	meaningless		
free dinners	increasingly		
free discussion	measured		
free download	increasingly		
free downloads	mediate		
free drink	increasingly		
free dumping	messianic		
free education	increasingly		
free education	militarized		
free election	increasingly		
free election	misaligned		
free elections	increasingly		
free elections	miserable		
free email	increasingly		
free employers	mobile		
free encyclopaedia	increasingly		
free energy	mobilized		
free enough	increasingly more		
free enterprise	increasingly		
free enterprise	moving		
free entry	increasingly		
free environment	multipolar		
free events	increasingly		
free exchange	multi-polar		
free exercise	increasingly		
free exhibition	murky		
free exhibitions	increasingly		
free expression	narrow		

free expression	increasingly		
free fall	nationalist		
free film	increasingly		
free financiers	negative		
free-fire zones	increasingly		
free flow	nervous		
free flowing	increasingly niche		
free food	increasingly noisy		
free food	increasingly		
free functioning	noted		
free gift	increasingly		
free glass	oblivious		
free hand	increasingly		
free hand	observe		
free handouts	increasingly		
free health	obsessive		
free healthcare	increasingly		
free healthcare	obsolete		
free hit	increasingly		
free import	obstructive		
free India	increasingly		
free individuals	obvious		
free information	increasingly		
freeing natural	online		
free inquiry			
free internet			
free judgment			
free judiciary			
free kerosene			
free kick			
free kippers			

free labour			
free land			
free lands			
free law			
free leaflet			
free library			
free life			
free loader			
free love			
free lunch			
free lunch			
freely abandon			
freely accessible			
freely admits			
freely agree			
freely associated			
freely available			
freely borrowing			
freely choose			
freely choose			
freely chose			
freely chosen			
freely confessed			
freely copying			
freely deciding			
freely deliberating			
freely describes			
freely discussed			
freely document			
freely elected			
freely held			

freely leaked			
freely link			
freely made			
freely mixing			
freely offered			
freely opt			
freely pick			
freely practice			
freely purchase			
freely reported			
freely revealed			
freely scattering			
freely selling			
freely singled			
freely sold			
freely speaks			
freely spread			
freely sprinkled			
freely strut			
freely traded			
freely transferable			
freely travel			
freely traverse			
free man			
free manner			
free market			
free market			
free-market			
capitalism			
free-market			
capitalism			

free marketers			
free marketer			
free-market			
fundamentalism			
free markets			
free markets			
free meal			
free meals			
free media			
free mediation			
free megabytes			
free men			
free midday			

Appendix H List of English stop words

English stop words							
A	At	ever	however	me	other	their	what
able	be	every	i	might	our	them	when
about	because	for	if	most	own	then	where
across	been	from	in	must	rather	there	which
after	but	get	into	my	said	these	while
all	by	got	is	neither	say	they	who
almost	can	had	it	no	says	this	whom
also	cannot	has	its	nor	she	'tis	why
am	could	have	just	not	should	to	will
among	dear	he	least	of	since	too	with
an	did	her	let	off	so	us	would
and	do	hers	like	often	some	wants	yet
any	does	him	likely	on	than	was	you
are	either	his	may	only	that	we	your
as	else	how		or	the	were	