IREvent2Story: A Novel Mediation Ontology and Narrative Generation

VenuMadhav Kattagoni
Center for Exact Humanities
IIIT Hyderabad
venumadhav.katagoni@gmail.com

Navjyoti Singh
Center for Exact Humanities
IIIT Hyderabad
singh.navjyoti@gmail.com

Abstract

Event detection is a key aspect of story development which is composed of multiple narrative layers. Most of the narratives are template-based and follow a narration theory. In this paper, we demonstrate a narrative from events detected in the international relations domain along with classification of events using our novel mediation ontology. We also introduce a novel method of classifying events through the mediation ontology. Our methodology involves action classification based on the verb categorization of Beth Levin, its arguments determined by universal dependencies and word2vec. The selected feature space is a result of mapping language entities to ontological entities where we obtain substantially good results. The narration also presents interactions of international actors over various topics as well as other visualizations.

1 Introduction

[GP12] states that "The field of international relations concerns the relationships among the world's governments. But these relationships cannot be understood in isolation. They are closely connected with other actors (such as international organizations, multinational corporations, and individuals); with other social structures and processes (including economics, culture, and domestic politics); and with geographical and historical influences. These elements together power the central trend in international relations today — globalization". Actors [Kan09] in international relations include individuals, groups (including ephemeral groups like crowds), organizations(including corporate entities, both public and private) and all generally recognized countries (including states and related territories). Classification of the events detected is important so as to analyze the group as a whole rather than each event discretely. Thus, we propose a new Mediation Ontology for international relations. This new mediation ontology provides a correlation between language entities and ontological entities which is used for classification of events into the proposed categories. We use its result as one of the visualizations in the narration.

We present a brief background on event ontologies and event coding in conjunction with media in Section 2. Our new Mediation Ontology is presented in Section 3. We then write about our dataset in Section 4. In Section 5, our methodology and results for classification of events to identify the event type and description of

In: A. Jorge, R. Campos, A. Jatowt, S. Nunes (eds.): Proceedings of the Text2StoryIR'18 Workshop, Grenoble, France, 26-March-2018, published at http://ceur-ws.org

the features used along with machine learning techniques for classification are discussed. In subsequent sections, narratives and visualizations are demonstrated in section 6 with a case study. We conclude the paper in section 7 with proposals on the future work sparked by this study.

2 Related Work

The last few decades have witnessed a considerable escalation in studies which are directed at event coding ontologies in the political domain. This kind of research began during the 1970s with the purpose of forecasting international conflict under the sponsorship of the U.S. Department of Defense Advanced Research Projects Agency (DARPA) [CR78], [AH88]. The kind of research that has been focused is mainly on:

1. the political event data coding ontologies.
2. the generation of the political event data.
3. forecasting of international conflict.

Our focus in this paper is restricted to international relations event coding ontology i.e., Ontology for international relations events or mediation types. Such ontologies include WEIS [Gol92], COPDAB [Aza80], CAMEO [GSYAJ02], IDEA [BBO$^+$03] etc. The WEIS Ontology is made up of 22 top-level categories that encompass actions such as Request or Grant. Each of these 22 top-level categories contain single level children which are further fine-grained. For example, the code 07 is the top-level code for reward with the sub-code 072 representing extended military assistance. The CAMEO ontology is an upgraded version of WEIS with mediation event types added to it. It is more fine-grained with 20 top-level categories that encompass actions such as Make-Statement or Protest. Each of these 20 top-level categories contain finer-grained categories in a hierarchical manner. For example, the code 14 is the top-level code for Protest with the sub-code 141 representing a general demonstration or rally. Under the code 141 is code 1411 which codes demonstrate or rally for leadership change. Thus, as one moves down the hierarchy of CAMEO, it becomes concise. Based on one's need, CAMEO or any event data coding schemes can be evolved using a mix-and-match framework whereby a researcher could adopt most of his or her coding categories from a standard set, and then elaborate on a smaller number of newer categories. Event coding using CAMEO [GSYAJ02] involves event detection and classification based on pattern matching from a large set of verb patterns, actors, compound nouns, compound verb phrases, reference to pronouns and deep parsing of sentences in news articles. Prior to CAMEO [GSYAJ02], event encoding was done manually based on the rules mentioned in the corresponding codebooks. Our work presented in this paper carves a similar problem by computing event types and narrative generation of the international events.

3 Mediation Ontology

Bercovitch [Ber97] defines mediation as "a process of conflict management, related to but distinct from the parties' own negotiations, where those in conflict seek the assistance of, or accept an offer of help from, an outsider (whether an individual, an organization, a group, or a state) to change their perceptions or behaviour, and do so without resorting to physical force or invoking the authority of law." He also mentions, "Mediation may well be the closest thing we have to an effective technique for dealing with conflicts in the twenty-first century". The main goals of this research are event classification and narrative generation so as to help journalists and researchers identify interactions among actors during international conflicts.

Our mediation ontology is inspired from social agents grouped under "The ontologies of Persons" mentioned by [Bic12]. He mentions, "Persons, so I argue, are special kinds of agents that arise in and are constituted in interactions with social and cultural processes, including other social persons, and thereby co-constitute the emergence base for those social and cultural realities." He adds that "In this model, agents are constituted in their interactive dynamics; such interactive dynamics is their ontology." Our mediation ontology is inspired from this concept of person and interactive dynamics. He also states that "Agents that develop to become participants, thus constitutive participants, in those environments are themselves, therefore, emergent kinds of agents — social agents:persons." A person's interactive dynamics can be inferred from the pronouncements he makes, the engagements he has with other persons, his response to another person's opinions and the use of force in unhealthy relations. In a similar manner, an actor in international relations interacts with other actors through pronouncements, engagements, responses and force mechanisms. The responses and force mechanisms of an actor determine the pronouncements and engagements made by peers. This is because pronouncements and

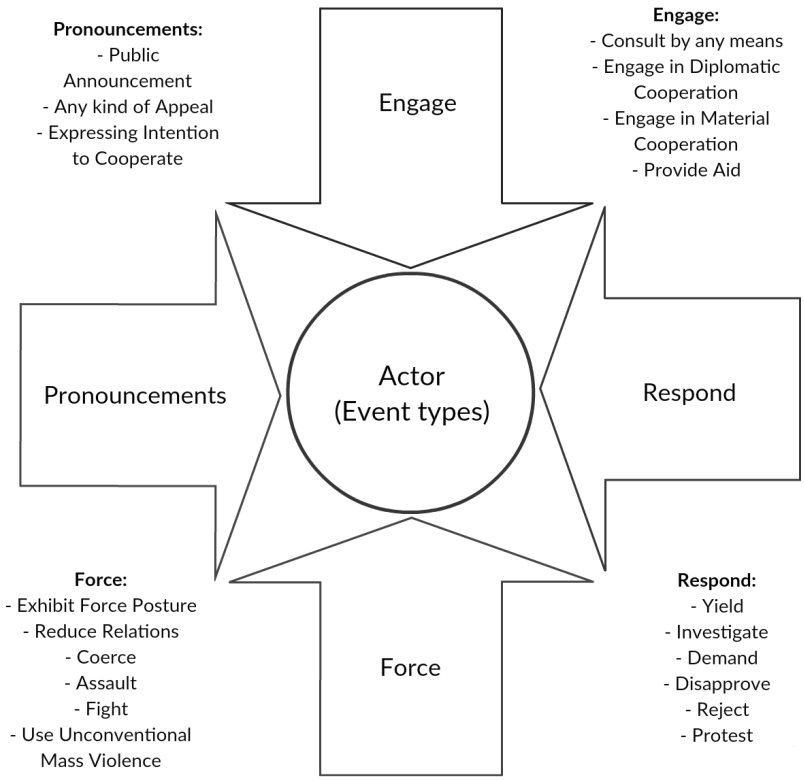


Figure 1: Mediation Ontology

engagements happen only when some kind of base event has occurred. Hence, force mechanisms and responses are ground event types whereas pronouncements and engagements are lateral event types. As mentioned by [GP12], the relations among the actors is what the field of international relations is concerned about. Therefore, multiple actors coming together would determine international relations.

The motivation behind a new mediation ontology arises from the 250+ classes in CAMEO [GSYAJ02]. There is an overlap in the mappings from verbs to classes in their verb dictionaries. Also, verb classification is an extremely context-sensitive exercise. Hence, we map language entities with ontological entities while proposing a new statistical model of event classification system which meets all our requirements. We classified an event type into four classes instead of the 20 top-level classes that CAMEO [GSYAJ02] consists of (with nearly 250+ subclasses). Since CAMEO [GSYAJ02] is widely used, we mapped the CAMEO [GSYAJ02] categories as following in order to come-up with the current definitions of event types.

1. Pronouncements
   - Declining to comment, making pessimistic and optimistic comment, claiming, denying, empathetic, accord, symbolic act, policy option.
   - Appeal for material or diplomatic cooperation, aid, political reform, negotiation, settling disputes, accepting mediation.
   - Expressing intent to cooperate, material or diplomatic cooperation, providing aid, political reform, yield, negotiating, settle disputes, mediation.
   - CAMEO Classes - 01, 02, 03.
2. Engage
   - Consult, discuss, meet, negotiate, mediate.
   - Engaging in diplomatic, material, economic, military, judicial, intelligence cooperation, endorse, defend verbally, support, recognize, apologize, forgive, formal agreement.

Code	Class Name
01	Make Public System
02	Appeal
03	Express intent to cooperate
04	Consult
05	Engage in Diplomatic Cooperation
06	Engage in Material Cooperation
07	Provide Aid
08	Yield
09	Investigate
10	Demand
11	Disapprove
12	Reject
13	Threaten
14	Protest
15	Exhibit Force Posture
16	Reduce Relation
17	Coerce
18	Assault
19	Fight
20	Use unconventional mass violence

Table 1: CAMEO's top-level classification

Class	Training Data	Testing Data	Total
Pronouncements	38130	4237	42367
Engage	23136	2571	25707
Respond	16275	1809	18084
Force	15510	1724	17234
Total	93051	10341	103392

Table 2: Dataset Description

- CAMEO Classes - 04, 05, 06, 07.

3. Respond

   - Any type of response in the form of yield, investigate, demand, disapprove, reject, threaten, protest.
   - CAMEO Classes - 08, 09, 10, 11, 12, 13, 14

4. Force

   - Any type of force posture, reducing relations, coerce, assault, fight, use unconventional mass violence.
   - CAMEO Classes - 15, 16, 17, 18, 19, 20

Our mediation ontology is described in figure 1. All the 20 top-level CAMEO classes are described in table 1.

4 Dataset

Our system listens to 248 media feeds[1] for news articles daily. We use the methodology proposed in [KS18] to extract events between September 1, 2017 and September 30, 2017. Since we mapped our categorical information with CAMEO, we used the Petrarch system [CN17] based on CAMEO [GSYAJ02] to generate its event types and map to our categories. We generated a total of 103392 events distributed across all the four classes. Detailed description regarding our dataset is described in table 2.

[1] http://ceh.iiit.ac.in/international_relations/source.txt

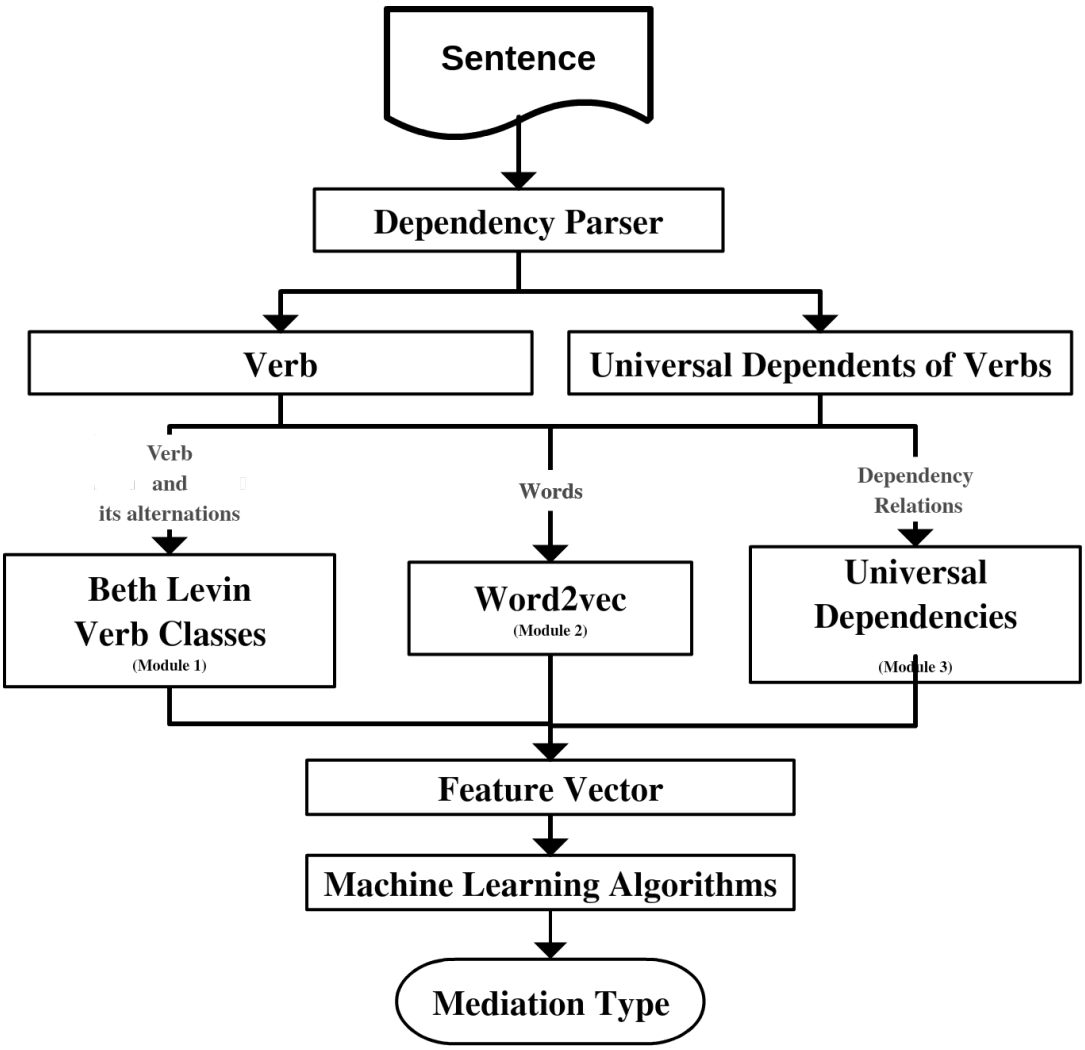


Figure 2: Methodology.

5 Methodology and Results

The methodology is described in figure 2. The sentence in which the event is detected is sent to Stanford Dependency parser[CM14] to identify the action verb and its dependencies. The identified action verb, its arguments and universal dependency relations [NdMG$^+$16] are passed through 3 different modules which finally unite to form our feature space.

1. In the first module, we identify the class of the verb with respect to Beth Levin Verb Classes [Lev93] considering verb and its alternations in the sentence. We chose 59 classes which are relevant to our classification. Mapping between Beth Levin [Lev93] Verb Classes and our mediation ontology categories is described in table 3.
2. In the second module, the verb and its arguments which are found using the universal dependencies are converted to vectors using Google word2vec pretrained model [TMS]. All the argument vectors are added to the verb vector.
3. In the third module, all the universal dependency relations [NdMG$^+$16] of the verb with its arguments are taken into account. Since, there are 40 universal dependencies as mentioned in [NdMG$^+$16], we consider a 40 dimensional vector which is normalized by the total number of dependencies detected in the sentence.

The results of a few Machine Learning algorithms (viz., Logistic Regression, Random Forest, ensemble of Logistic Regression and Random Forest and Multi-Layer Perceptron) on the feature space obtained from the

Class	Beth Levin Verb Classes
Pronouncements	Characterize Verbs , Appeal Verbs , Long Verbs , Verbs of Transfer of a Message , Tell Verbs, Verbs of Manner of Speaking , Say Verbs , Complain Verbs , Reflexive Verbs of Appearance
Engage	Pit Verbs , Drive Verbs , Contribute Verbs , Verbs of Future Having , Verbs of Exchange , Build Verbs , Grow Verbs , Create Verbs , Performance Verbs , Dub Verbs , Conjecture Verbs , Admire Verbs , Judgment Verbs , Correspond Verbs , Meet Verbs , Talk Verbs , Chitchat Verbs , Dine Verbs , Gorge Verbs , Verbs of Spatial Configuration , Verbs of Contiguous Location , Verbs of Inherently Directed Motion , Roll Verbs , Verbs that are not Vehicle Names , Accompany Verbs
Respond	Banish Verbs , Manner Subclass , Verbs of Possessional Deprivation: Cheat Verbs , Get Verbs , Hold Verbs , Verbs of Concealment , Separate Verbs , Split Verbs , Disassemble Verbs , Amuse Verbs , Verbs of Assessment , Search Verbs , Investigate Verbs , Advise Verbs , Break Verbs , Bend Verbs , Other Alternating Verbs of Change of State , Verbs of Lingering
Force	Throw Verbs , Hit Verbs , Swat Verbs , Sight Verbs , Murder Verbs

Table 3: Mapping between Mediation categories and Beth Levin Classes.

Method (One Vs Rest Classifier)	Precision	Recall	F1-Score	Accuracy
Logistic Regression	0.79	0.80	0.79	0.78
Random Forest	0.75	0.77	0.76	0.73
Ensemble (Logistic Regression + Random Forest)	0.78	0.80	0.79	0.77
Multi-layer Perceptron	0.80	0.80	0.80	0.80

Table 4: Results

above methodology are described in the table 4. The ensemble technique we used is majority rule voting. All the hyper-parameters of Multi-Layer Perceptron [HM94] are described in table 5. All the metrics (precision, recall and accuracy) are the average of the corresponding class metrics. The optimum result was obtained using Multi-Layer Perceptron [HM94] with precision, recall and accuracy of 80%. The results are favourable using Multi-Layer Perceptron [HM94] because of the backpropagation training algorithm. It is worth to note that all the other Machine Learning algorithms produce nearly same results which gives a strong base for the choice of our feature space.

6 Narrative Generation

We use the event model presented by [KS18] for extracting events with attributes date-time, location, actors, media-source, event-title, source-url, sentence. Extending this model with action (verb) and action-type (event-type), our mediation ontology adds two new attributes — action and action-type. These attributes capture individual actions when events are grouped with topics helping in capturing subtler details of the generated narrative. Updated event model is shown in figure 3.

Our system visualizes the event actor interaction using graphical, topical, geographical and temporal features. These visualizations help journalists and researchers in IR domain in understanding the interactions among actors.

Hyper-Parameter	Congifured Description
Hidden layer sizes	(100,)
Activation function for hidden layer	Rectified linear unit function
Solver	adam (Stochastic gradient-based optimizer proposed by [KB14])
Learning Rate	Constant with an initial learning rate of 0.001
Maximum number of iterations until solver converges	200
Tolerance for optimization	0.001

Table 5: Hyper-parameters of Multi-Layer Perceptron

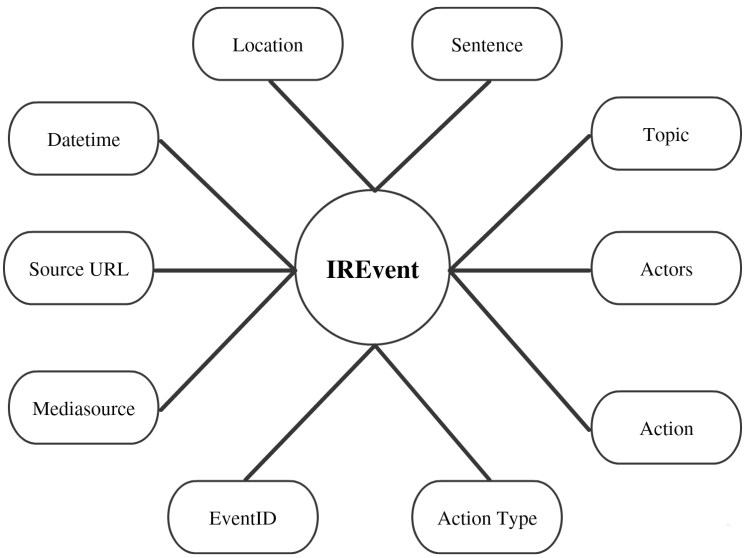


Figure 3: Updated Event Model with action and action-type attributes added.

We consider the definition of narrative event chain by [CJ08]. [CJ08] states that "A narrative event is a tuple of an event and its participants, represented as typed dependencies. A narrative event chain is a partially ordered set of narrative events that share a common actor (protagonist)." Our narrative event chain is also a partially ordered set of narrative events that shares two common actors.

The narrative starts with the description of two actors followed by various visualizations describing the interaction of the actors.

1. **Bar Chart:** The bar chart depicts frequency of the events reported involving both the actors with the most reported topics of interaction and unique events.

2. **Line Chart:** The line chart compares the frequency of the events reported involving individual and combined actors. This provides how important the interaction of the mentioned actors is in comparison to total number of interactions reported with all other actors.

3. **Mediation Class Events:** The mediation class events gives the number of events reported in each category of the mediation ontology. This visualization provides comparison among the classes of interaction.

4. **Graph Visualization:** The graphical visualization represents the interaction wherein nodes are the actors and the edges are topics of interaction. This visualization helps place an actor level context to the conflict. Hovering on the nodes and edges provides with the actor name and the topics of interaction respectively.

5. **Topic Cloud:** The topical visualization helps situate the gravity of the topics spoken of and thus giving a subjective view of the conflict. It is created based on the frequency of the words used in the topic.

6. **Geographic Visualization:** The geographical visualization brings to the fore the narrative about actor's stakes in the international conflict and thus add a geopolitical persona to the event. The importance of the actor increases with the increase in the number of sub-actors. Hovering on the circle gives all the actors and sub-actors involved in that area.

7. **Timeline Visualization:** The timeline visualization helps bring a coherency to the event-actor duo and places the interaction over a span of the dialogue until its closure.

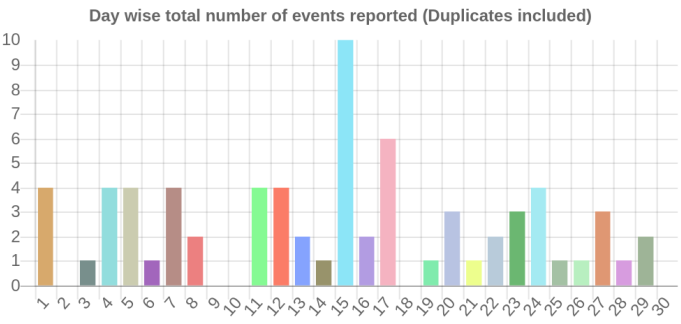
We show an example of a narrative involving Japan and South Korea as common actors surrounded by a narrative event chain. All the narrative visualizations between them are shown in figure 4.

A live prototype of the system is available here: `http://ceh.iiit.ac.in/international_politics/`

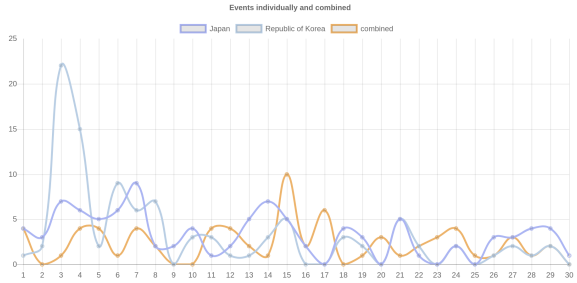
Major events

Major events between the two actors are: **RECENT TEST SECOND MISSILE NORTH KOREA SOUTH KOREAN JAPANESE OFFICIALS**

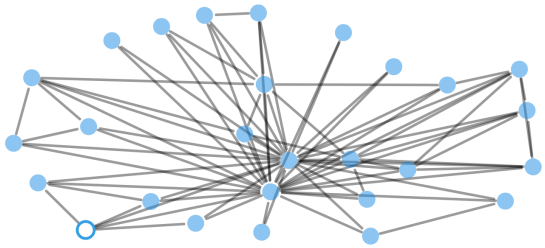
Total number of unique event occurence: **28**



(a) Bar Chart

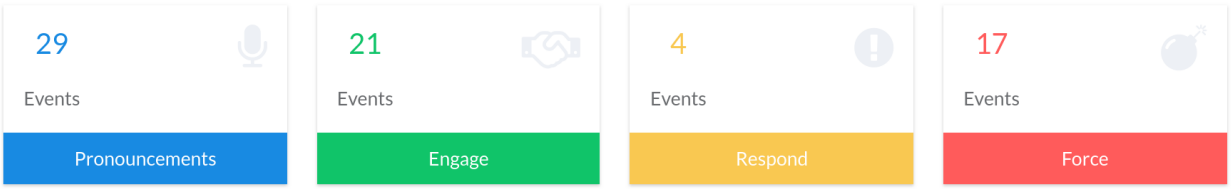


(b) Line Chart



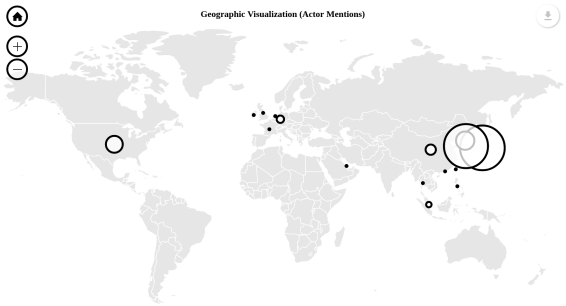
(c) Graph Visualization

Mediation Class wise events



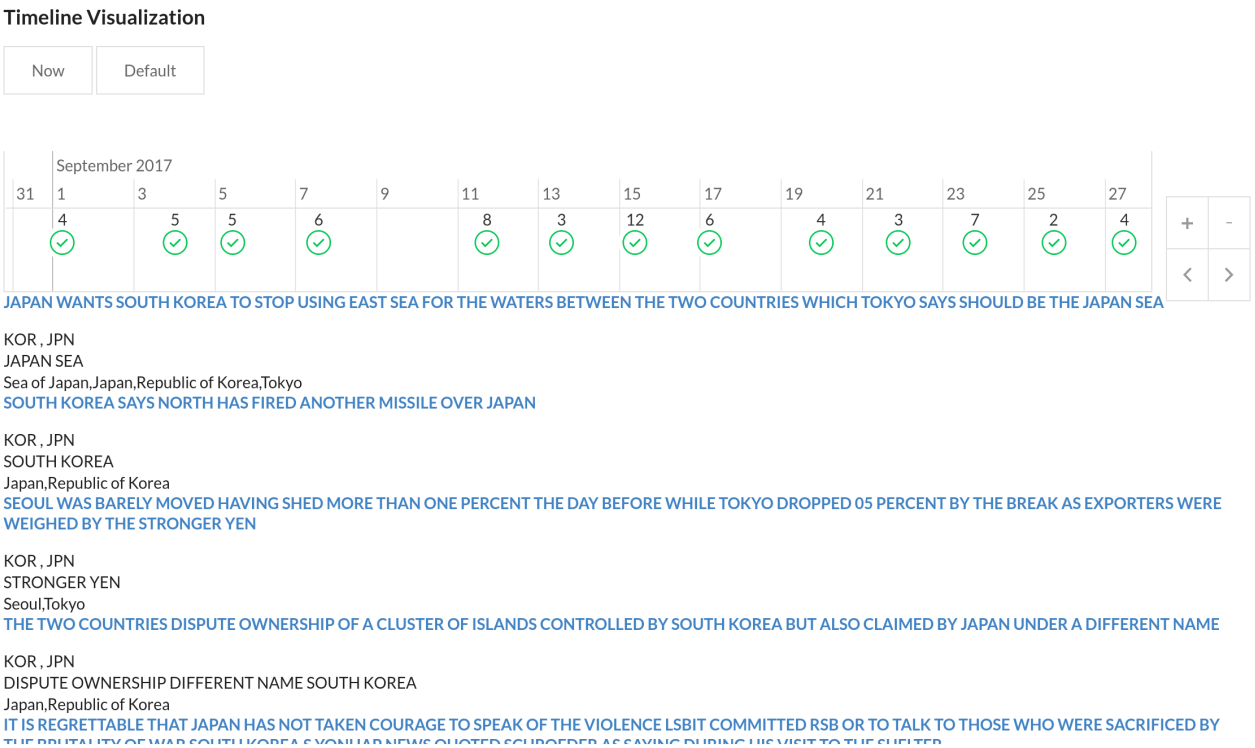
(d) Mediation Class Counts

(e) Topic Visualization



(f) Map Visualization

Figure 4: All visualizations in the narrative.



(g) Timeline Visualization

Figure 4: All visualizations in the narrative (cont.)

7 Conclusion and Future Work

Our paper described a novel ontology for categorization of the news events which helps in event classification in the international relations. We built a system, IREvent2Story that helps identify the various narrative features behind the events. Our ontology is a step towards framing a further attuned vocabulary for discussion of any international exchange thus setting the base for more theory work on ideating a framework for mapping not only political entities but also include non-political entities in a similar framework. Our ontology helps to not only drive meaning through the vast news data corpus but also acts as a step towards conceptualizing self-hydrating and sustaining systems of data journalism that will usher with the Web 2.0. The vision of this research is also to find the parts in the news article where the conflict is framed and compare this conflict among various news sources.

References

[AH88] Stephen J Andriole and Gerald W Hopple. *Defense Applications of Artificial Intelligence*. Lexington Books, 1988.

[Aza80] Edward E. Azar. The conflict and peace data bank (copdab) project. *Journal of Conflict Resolution*, 24(1):143–152, 1980.

[BBO+03] Doug Bond, Joe Bond, Churl Oh, J. Craig Jenkins, and Charles Lewis Taylor. Integrated data for events analysis (idea): An event typology for automated events data development. *Journal of Peace Research*, 40(6):733–745, 2003.

[Ber97] Jacob Bercovitch. Mediation in international conflict: An overview of theory, a review of practice. *Peacemaking in international conflict: Methods and techniques*, pages 125–154, 1997.

[Bic12] Mark H. Bickhard. *The emergent ontology of persons*, page 165180. Cambridge University Press, 2012.

[CJ08] Nathanael Chambers and Dan Jurafsky. Unsupervised learning of narrative event chains. *Proceedings of ACL-08: HLT*, pages 789–797, 2008.

[CM14] Danqi Chen and Christopher Manning. A fast and accurate dependency parser using neural networks. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 740–750. Association for Computational Linguistics, 2014.

[CN17] John Beieler Clayton Norris, Philip Schrodt. Petrarch2: Another event coding program. *Journal of Open Source Software*, 2017.

[CR78] Nazli Choucri and Thomas W Robinson. *Forecasting in international relations: Theory, methods, problems, prospects*. Freeman, 1978.

[Gol92] Joshua S. Goldstein. A conflict-cooperation scale for weis events data. *Journal of Conflict Resolution*, 36(2):369–385, 1992.

[GP12] Joshua S Goldstein and Jon C Pevehouse. *International relations*. Pearson/Longman, 10 edition, 2012.

[GSYAJ02] Deborah J Gerner, Philip A Schrodt, Omür Yilmaz, and Rajaa Abu-Jabr. Conflict and mediation event observations (cameo): A new event data framework for the analysis of foreign policy interactions. *International Studies Association, New Orleans*, 2002.

[HM94] Martin T Hagan and Mohammad B Menhaj. Training feedforward networks with the marquardt algorithm. *IEEE transactions on Neural Networks*, 5(6):989–993, 1994.

[Kan09] H. Kan. *In Government and Politics*, volume II, chapter Actors in World Politics. UNESCO-EOLSS, 2009. edited by Masashi Sekiguchi, Tokyo Metropolitan University, Japan.

[KB14] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014.

[KS18] VenuMadhav Kattagoni and Navjyoti Singh. Towards an unsupervised learning method to generate international political event data with spatiotemporal annotations. In Andrew U. Frank, Christine Ivanovic, Francesco Mambrini, Marco Passarotti, and Caroline Sporleder, editors, *Proceedings of the Second Workshop on Corpus-Based Research in the Humanities CRH-2*, volume 1 of *Gerastree proceedings*, pages 105–112, 2018.

[Lev93] B. Levin. *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press, Chicago, IL, 1993.

[NdMG+16] Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajic, Christopher D Manning, Ryan T McDonald, Slav Petrov, Sampo Pyysalo, Natalia Silveira, et al. Universal dependencies v1: A multilingual treebank collection. In *LREC*, 2016.

[TMS] Quoc V. Le Tomas Mikolov and Ilya Sutskever. word2vec.