IREvent2Story: A Novel Mediation Ontology and Narrative Generation

VenuMadhav Kattagoni and Navjyoti Singh
Center for Exact Humanities,
International Institute of Information Technology, Hyderabad, India
venumadhav.kattagoni@gmail.com, singh.navjyoti@gmail.com

Abstract

Event detection is a key aspect of story development which is itself 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 using our novel mediation ontology. We also introduce a novel method of classifying events through the mediation ontology using Beth Levin Verbs Classification, word2vec and Universal Dependencies. The selected feature space is a result of mapping language entities to ontological entities where we obtain substantially good results. Our methodology involves action classification based on the verb categorization of Beth Levin and its arguments determined by universal dependencies. The narration also presents interactions of international actors over various topics and other visualizations which would help journalists and researchers in the international relations domain.

Keywords: international relations, international politics, international news, ontology, event analysis, narrative generation, machine learning

1. Introduction

International relations are mostly framed by the pronouncements, engagements, responses, comments or force postures made by the actors. Actors (Kan, 2009) 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 discreetly. Thus, we propose a new Mediation Ontology for international relations. This new mediation ontology also provides a correlation between language entities and ontological entities which is used for classification of events into the proposed categories. We use its result in the narrative generation. Journalists often dig deep and push back against conventional wisdom, take time and resources from media companies - many of which are time-taking. We also intend to reduce this process of going through the history of similar articles from most of the media companies.

We present a brief background on the event ontologies and event coding in conjunction with media in Section 2. We present our new Mediation Ontology in Section 3. We then present our dataset in Section 4 and methodology and results for classification of events to identify the event type and also to describe the features used along with the machine learning techniques for classification in Section 5. In subsequent sections, narratives and visualizations demonstrated in section 6. We end 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) (Choucri and Robinson, 1978), (Andriole and Hopple, 1988). The kind of research that has been focused 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 relation event coding ontology i.e., Ontology for international relations events or mediation types. Such ontologies include WEIS (Goldstein, 1992), COPDAB (Azar, 1980), CAMEO (Gerner et al., 2002), IDEA (Bond et al., 2003) 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 contains single level children which are more 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 contains finer-grained categories in a hierarchical manner. For example, the code 14 is the top-level code for Protest with the sub-code 142 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 more fine-grained. 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.
Our work presented in this paper carves a similar problem by computing event types and narrative generation of the international events.

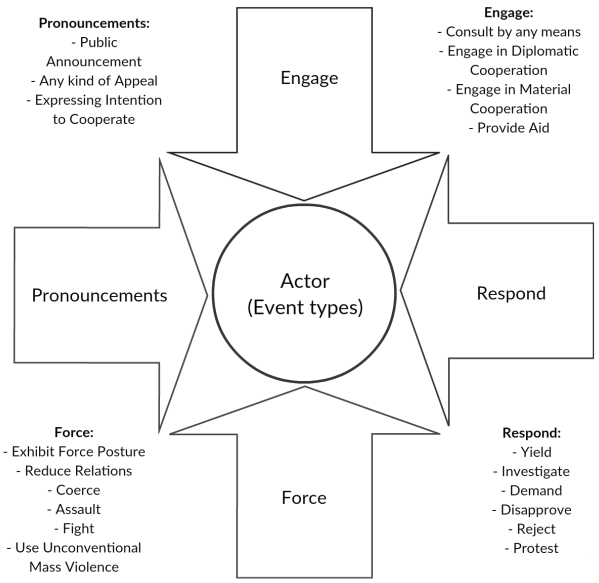

Figure 1: Mediation Ontology

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

3. Mediation Ontology

Bercovitch (Bercovitch, 1997) 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 behavior, 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 goal of this research is narration generation so as to help journalists and researchers identify interactions among actors during international conflicts. Since there are 250+ classes, there is overlap in their mappings from verbs to classes. 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 broadly four classes rather than 20 broad classes that CAMEO (Gerner et al., 2002) consists of (with nearly 250+ sub-classes).

This idea was inferred from the concept of a person. A person could be said to make some pronouncements, have certain engagements with other persons, respond to an another person's opinions and make 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 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. Therefore, multiple actors coming together would determine international relations. Our mediation ontology is described in figure 1. We mapped the CAMEO (Gerner et al., 2002) 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 dispute, 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.
   - 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

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

Class	Beth Levin Verb Classes
Pronouncements	Characterize Verbs , Appeal Verbs , Long Verbs , Verbs of Transfer of a Message , Tell , 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 , 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.

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

All the 20 broad CAMEO classes are described in table 1.

4. Dataset

Our system listens to 248 media feeds[1] for news articles daily. We used our previous work (Kattagoni and Singh, 2018) to extract events from the news articles dated between August 15, 2017 and September 30, 2017. Since we mapped our categorical information with CAMEO, we used the Petrarch system (Clayton Norris, 2017) based on CAMEO (Gerner et al., 2002) to generate data and map to our categories. The same data is fed to Petrarch (Clayton Norris, 2017) system. We generated a total of 103,392 events distributed across all the four class. Detailed description regarding dataset is described in table 2

5. Methodology and Results

The methodology is described in figure 2. The sentence in which the event is detected is sent to a Dependency parser to find verb and its dependencies. This passes 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 (Levin, 1993) considering verb and its alternations. We chose 59 classes which are relevant to our classification. Refer to Beth Levin (Levin, 1993) Verb Classes in table 3

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

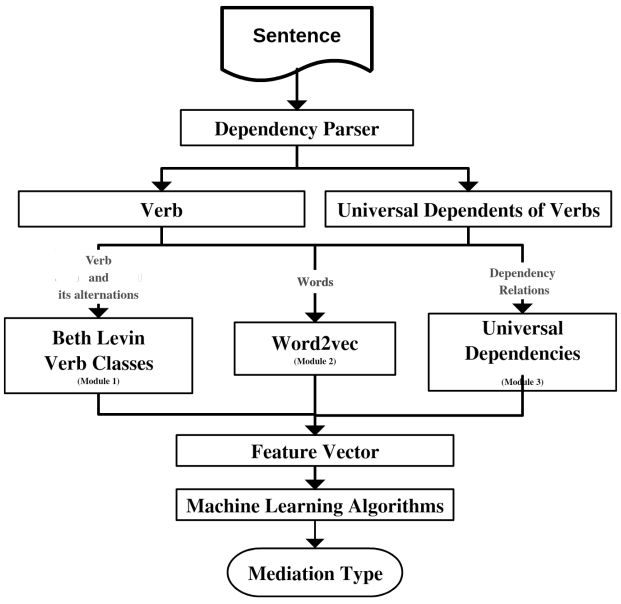

Figure 2: Methodology.

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 (Tomas Mikolov and Sutskever,). All the argument vectors are added with the verb vector.
3. In the third module, all the universal dependency relations (Nivre et al., 2016) of the verb with its arguments are taken into account.

The results of a few Machine Learning algorithms on the feature space obtained from the above methodology are described in the table 4. All the metrics (precision, recall and

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

accuracy) are the average of the corresponding class metrics. The optimum result was obtained using Multi-layer Perceptron (Hagan and Menhaj, 1994) with precision, recall and accuracy of 80%. The results are favourable using MLP 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 inferred event model from our previous work (Kattagoni and Singh, 2018) 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 types. These attributes capture individual actions when events are grouped with topics helping in capturing subtler details of the generated narrative. Our system visualizes the event actor interaction using graphical, topical, geographical and temporal features. The graphical visualization represents the interaction wherein nodes are the actors and its connected entities and the edges are topics. This visualization helps place an actor level context to the conflict. The topical visualization helps situate the gravity of the topics spoken of and thus giving a subjective view of the conflict. The geographical visualization helps corner the narrative about actor's stakes in the conflict and the geopolitical persona to the event. The temporal visualization helps bring a coherency to the event-actor duo and place the interaction over a span of the dialogue until its closure. A live prototype of the system is available here: `http://ceh.iiit.ac.in/international_politics/`

7. Conclusion and Future Work

Our paper described a novel ontology for categorization of the news corpus and help in event detection in the international fora. 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.

8. References

Andriole, S. J. and Hopple, G. W. (1988). *Defense Applications of Artificial Intelligence*. Lexington Books.

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

Bercovitch, J. (1997). Mediation in international conflict: An overview of theory, a review of practice. *Peacemaking in international conflict: Methods and techniques*, pages 125–154.

Bond, D., Bond, J., Oh, C., Jenkins, J. C., and Taylor, C. L. (2003). Integrated data for events analysis (idea): An event typology for automated events data development. *Journal of Peace Research*, 40(6):733–745.

Choucri, N. and Robinson, T. W. (1978). *Forecasting in international relations: Theory, methods, problems, prospects*. Freeman.

Clayton Norris, Philip Schrodt, J. B. (2017). Petrarch2: Another event coding program. *Journal of Open Source Software*.

Gerner, D. J., Abu-Jabr, R., Schrodt, P. A., and ÃmÃ¼r Yilmaz. (2002). Conflict and mediation event observations (cameo): A new event data framework for the analysis of foreign policy interactions. In *of Foreign Policy Interactions.â Paper presented at the International Studies Association*.

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

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

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

Kattagoni, V. and Singh, N. (2018). Towards an unsupervised learning method to generate international political event data with spatiotemporal annotations. In *the second edition Workshop on Corpus-Based Research in the Humanities (CRH)*.

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

Nivre, J., de Marneffe, M.-C., Ginter, F., Goldberg, Y., Hajic, J., Manning, C. D., McDonald, R. T., Petrov, S., Pyysalo, S., Silveira, N., et al. (2016). Universal dependencies v1: A multilingual treebank collection. In *LREC*.

Tomas Mikolov, Quoc V. Le and Ilya Sutskever.). *word2vec*.