Decoding Country Perspectives on "Rule of Law" in U.N. Security Council Debates: A Comparative Analysis Leveraging conText Embedding Regression

Jingyi Chen
jic090@ucsd.edu

GitHub Code Repository:
https://github.com/jchensd/POLI_179_Jingyi_Chen.git

1 Introduction

Rule of law is a fundamental concept in international conflicts, collaborations, and state governance. The mentioning of rule of law in the U.N. Security Council had increase in the last two decades. While it was not frequently brought up during the debates, with about 10 percents of speeches that mentioned it among all speeches per year in the past 20 years (Figure 1), the manner in which countries discuss this concept can provide implications about their values and goals on the international stage. In this project, by applying the à la carte (ALC) Embedding Khodak et al. (2018) and the conText Embedding Regression (Rodriguez, Spirling and Stewart 2023) to compare the contextual meanings of "rule of law" in the United Nations Security Council debates between the U.S. and China and across countries with different democracy scores (Economist Intelligence Unit 2010), I tried to understand whether there were distinctions in how countries viewed this concept.

2 Related Work

In the analysis on Congressional debates, Rodriguez, Spirling and Stewart (2023) examined how the contextual usage of words by Democrats and Republicans systematically differed by applying the ALC embedding and the conText regression model. They found that "immigration" was the most partisan word, which Democrats tended to associated with "reform" and Republicans tended to associate with "enforce" and "regularize".

There have been relatively limited literature applying computational text analysis methods to study how countries differed in their interpretations of certain concepts in the United Nations Security Council debates. Sakamoto (2019) applied the word embedding method to study to what degree countries shared perceptions of "threat to the peace". The findings suggest that the major powers in the Council, such as United States and Russia, have significantly similar perceptions of potential threats to international peace and security.

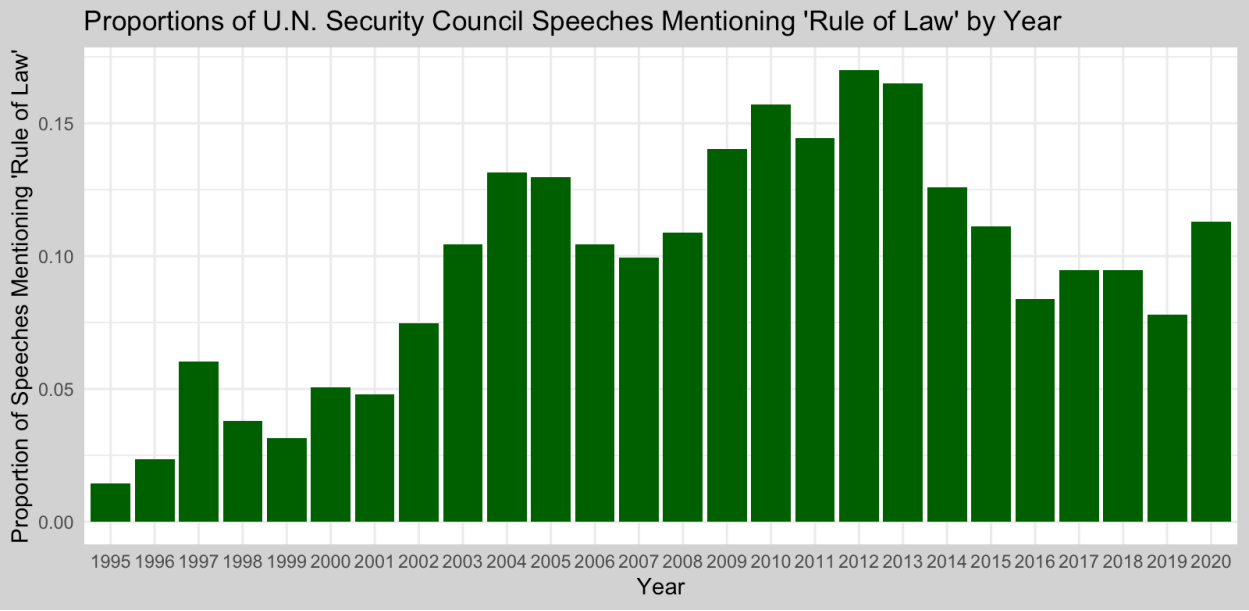

Figure 1: Proportions of U.N. Security Council Speeches Mentioning ”Rule of Law” by Year

3 Data

3.1 Data 1

The main data analyzed in this project is the United Nations Security Council debates (Schoenfeld et al. 2019). The filtered data include 9441 speeches that mentioned ”rule of law” from January 1995 to December 2020 and their document IDs, as well as the metadata including country, speaker, and date. I merged the separate files of the original data package using the shared document IDs. Because the primary objective of this project is to examine how the phrase ”rule of law”, or the targeted/focal term, is used by different countries, it was made into one term, i.e. ”ruleoflaw”. Also, because I am also interested in whether some countries are more likely to link ”human rights” to ”ruleoflaw”, the former phrase is also made into one term, i.e. ”humanrights”. ”United Nations” and ”Security Council” were also made into single terms. The whole dataset of speeches were filtered into only those that mentioned ”ruleoflaw”.

Then, I converted the column of texts into a corpus and tokens. I also conducted common text pre-processing using the ”quanteda” package, including removing punctuation, symbols, numbers, separators, English stopwords, words with 2 or fewer characters and words appearing less than 5 times in the corpus (I used ”padding=TRUE” to leave an empty string where the removed tokens previously existed), and converting words to lower case.

3.2 Data 2

The Democracy Index is a value indicating a country's level of democracy, ranging with 0 to 10 (Economist Intelligence Unit 2010). I used democracy index data in 2010 to analyze the difference in the meaning of rule of law among countries with higher and lower democracy scores.

4 Methodology

4.1 Method 1: à la carte (ALC) Embedding

I used the "conText" package (Rodríguez 2023) developed by Rodriguez, Spirling and Stewart (2023) to conduct analysis in R. In order to examine the contextual meanings of "ruleoflaw" by analyzing words surrounding this term, I collected 6 words before and after every instance in which "ruleoflaw" was mentioned among the speeches as well as the metadata associated with the instance. A document feature matrix was built, where the rows were the total 16145 instances of "ruleoflaw". Then, I estimated GloVe embeddings from the 9441 U.N. Security Council speeches using the "text2vec" package. A transform matrix (Matrix A) was computed using the "compute-transform" function. Matrix A helps weight the common, uninformative words lower in the embeddings.

The ALC embedding method estimates the embedding of the targeted word in a specific context by taking the average of the embeddings of words surrounding this word (e.g. 6 words before and after the targeted word) (Khodak et al. 2018). I obtained the embeddings of the words surrounding "ruleoflaw" from the GloVe model and applied Matrix A to construct a document-embedding matrix. With the GloVe embedding model, I found nearest neighbor features for country-specific "ruleoflaw" embeddings.

I also computed cosine similarities between China and U.S.'s "ruleoflaw" embeddings and available features in the corpus to find words that distinguish two countries the most. To obtain the ratios, it first computed the cosine similarities between the embeddings of "ruleoflaw" and available features in the corpus for China and the U.S. Then, it takes the ratio of cosine similarities between the two countries for each features.

4.2 Method 2: conText Embedding Regression

Using the conText regression model, I regressed the embedding of "ruleoflaw" over the 2010 Democracy Index. Before doing so, I merged the U.N. Security Council dataset with the 2010 Democracy Index data. Speeches by NGOs, supranational unions, and a few countries for which Democracy Index data were unavailable were removed from the dataset. There are 7879 speeches by countries left. After constructing the regression, I obtained the embeddings for countries at every 5th percentile of democracy index that were estimated by the regression and found nearest neighbor features. I mainly compared countries at 5th

Table 1: Nearest Neighbor Features for the Overall "ruleoflaw" Embedding

feature	rank	value
ruleoflaw	1	0.9355679
strengthening	2	0.8397012
respect	3	0.7969343
essential	4	0.7895955
governance	5	0.7894433
institutions	6	0.7815149
humanrights	7	0.7773125
promoting	8	0.7600160
ensuring	9	0.7440357
justice	10	0.7420294

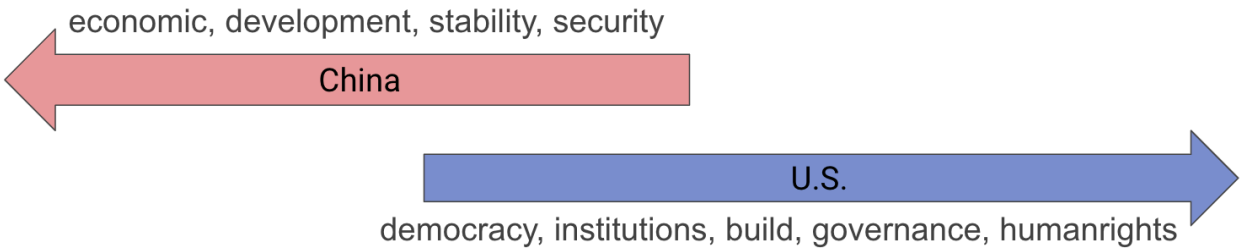

Figure 2: Ratios of Cosine Similarities (China vs. U.S.)

and 95th percentiles of democracy index by computing ratios of cosine similarities.

5 Results & Findings

5.1 Nearest Neighbors Based on ALC Embedding

Table 1 shows the nearest neighbors for the overall "ruleoflaw" embedding from the whole corpus (i.e. 9441 U.N. Security Council speeches that mentioned "ruleoflaw"). Overall, countries, NGOs, and supranational organizations emphasized the importance of rule of law and linked it to democratic concepts and human rights.

Table 2 presents the nearest neighbor features for the specific contexts of China and the U.S. Figure 2 displays the ratios of cosine similarities. The features with values larger than 1 are more associated with "ruleoflaw" in China's context compared to the U.S.'s context; vise versa, features with values smaller than 1 are more associated with "ruleoflaw" in the U.S.'s context compared to China's context. The interpretation is that the U.S. tends to understand rule of law as protection of individuals' rights during international crises and the institutions and governance in the international decision-making process. In contrast, China tends to link it to stability, security, and economic development.

Table 2: Nearest Neighbor Features for the ”ruleoflaw” Embedding within the China’s (left) and US’s (right) Context

feature	rank	value
ruleoflaw	1	0.7990605
stability	2	0.7661675
development	3	0.7271184
strengthening	4	0.7197165
institutions	5	0.7150364
essential	6	0.7013694
promote	7	0.6952098
security	8	0.6859930
efforts	9	0.6810040
strengthen	10	0.6791745

feature	rank	value
ruleoflaw	1	0.8580669
institutions	2	0.7776769
governance	3	0.7363304
strengthen	4	0.7262979
strengthening	5	0.7179879
respect	6	0.7093298
promote	7	0.6938324
essential	8	0.6928845
humanrights	9	0.6892298
build	10	0.6694275

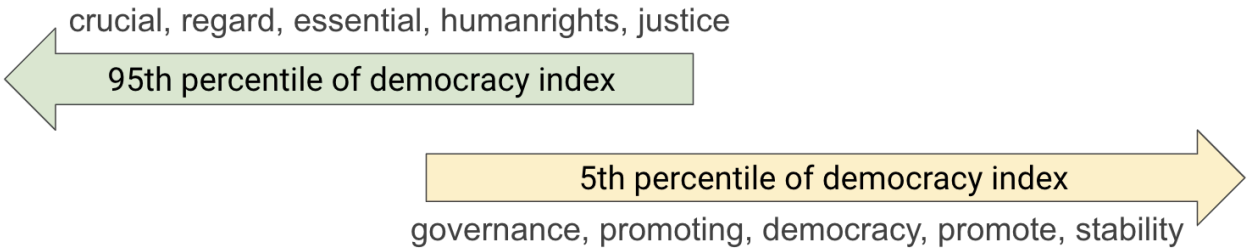

Figure 3: Ratios of Cosine Similarities (95th vs. 5th Percentitles of Democracy Index)

5.2 conText Embedding Regression Outputs

The regression model outputs a p-value of 0 and lower and higher confidence intervals of 0.06375997 and 0.07463132, indicating the regression result is significant. Based on the nearest neighbor features (not presented here due to page limitation) for embedding of ”ruleoflaw” for countries at every 5th percentile, as estimated by the regression, even countries at 95th percentile and 5th percentile have similar top nearest neighbors like ”strengthening” and ”respect”, indicating their shared acknowledgement of the importance of rule of law. However, when taking the ratios of cosine similarities, the results show that countries at 5th percentile tend to link the phrase to ”stability” (ratio: 0.9238250) and ”democracy” (0.9656132, not as significant as ”stability”), while those at 95th percentile tend to emphasize its importance and associate it with human rights (1.0519852) and justice (1.0468595). Figure 3 displays the ratios of cosine similarities between countries at 95th percentile and 5th percentile of democracy index.

Considering that China typically connects rule of law to stability, I removed all speeches by China from the dataset and ran the regression again. Still, ”stability” has the lowest cosine similarity ratio (0.9450454), suggesting that linking rule of law to stability is a common observation for countries with low democracy scores, not uniquely for China, though it is a very typical example.

6 Discussions & Conclusions

Based on the results from ALC embedding and conText Regression, this study found that, in the scenario of U.N. Security Council debates, countries like China and the U.S. showed different understandings of rule of law. The U.S. presents a traditional Western liberal-democratic interpretation of rule of law, while China possibly considers it as a pragmatic means to broader socioeconomic and political goals. Also, states with lower democracy scores tend to link rule of law to stability, which might suggest maintaining stability holds high importance for non-democratic countries.

One limitation of this study is the lack of understanding of why the countries associate rule of law with distinct words in the broader context of international politics. In the future, it would be helpful to integrate background insights about state policies and international relations to better understand what the findings suggest about countries' different perspectives regarding rule of law, and how this could help explain international conflicts and collaborations.

References

Economist Intelligence Unit. 2010. "Democracy Index." [Link]

Khodak, Mikhail, Nikunj Saunshi, Yingyu Liang, Tengyu Ma, Brandon Stewart, and Sanjeev Arora. 2018. "A La Carte Embedding: Cheap but Effective Induction of Semantic Feature Vectors." In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Melbourne, Australia Association for Computational Linguistics. [Link]

Rodriguez, Pedro L., Arthur Spirling, and Brandon M. Stewart. 2023. "Embedding Regression: Models for Context-Specific Description and Inference." *American Political Science Review* 117(4), p. 1255–1274. [Link]

Rodríguez, Pedro L. 2023. "conText - Quick Start Guide." Aug. [Link]

Sakamoto, Takuto. 2019. "On 'Threat to the Peace': Computational Analysis of the Speech Records of the United Nations Security Council." *Riron to hōhō* 34(2): 262–

Schoenfeld, Mirco, Steffen Eckhard, Ronny Patz, Hilde van Meegdenburg, and Antonio Pires. 2019. "The UN Security Council Debates." [Link]