Automated Extraction of a Game Model using Natural Language Processing: National Strategies of the Asian Infrastructure Investment Bank

Marin Nagase, Yu Maemura, Masahide Horita
The University of Tokyo, Tokyo, Japan
mnagase@ken-mgt.t.u-tokyo.ac.jp, maemura@civil.t.u-tokyo.ac.jp, horita@k.u-tokyo.ac.jp

Abstract— **This article presents a method for representing highly strategic events using natural language processing. In 2013, China proposed the creation of the Asian Infrastructure Investment Bank (AIIB) as the country continues to capitalize on its economic growth and makes moves to gain influence in regional development. International society is facing a big shift in power balances, where the existing sphere of influence wielded by financial institutions led by the US and Europe are challenged by new orders being established by the developing countries. The method proposed in this article automates the analysis of written material so that the strategies taken by people or organizations involved can be modelled. The method uses clue expressions and supervised machine learning to identify and extract sentences containing strategic relations from a corpus. The accuracy of extraction, defined as the sentences including strategic relations in the sentences that included clue expressions, was 60%. Coverage was calculated as the ratio of sentences correctly extracted in the sentences containing strategic relations within the sentences including clue expressions, and was 67%. In addition, a method for modelling the strategies extracted is proposed and the relations of strategies contained within news articles are represented in a strategic map. Strategic maps of a complex initiative such as the AIIB has revealed the links between various events that intuitively may be considered irrelevant, such as the AIIB and the annexation of Crimea. Furthermore, the change in game structure before and after UK participation in the AIIB became visible as the consequences of UK participation onto other players' preferences has been revealed.**

I. INTRODUCTION

In October 2015, China established the Asian Infrastructure Investment Bank (AIIB), a new multi-lateral development bank to fill in the funding gap created by the increasing demand for infrastructure in developing countries. The emergence of the AIIB may become a threat to the already existing financial institutions led by the United States and Europe, with the Chinese yuan replacing the US dollar as the world's premier reserve currency [1]. President Obama of the US remarked on April 17, 2015 that "If we do not help to shape the rules so that our businesses and our workers can compete in those markets, then China will set up rules that advantage Chinese workers and Chinese businesses [2]." As China gains influence through initiatives like the AIIB and the one belt, one road initiative, the Trans-Pacific Partnership (TPP) initiative can be considered not just a free trade pact, but a strategy for established economies to deal with China [1]. As we are witnessing rapid developments in the strategic actions of global economic players attempting to maximize their interdependent interests, the need for tools to analyze, understand, and predict the consequences of such strategies are increasingly apparent. In addition, the amount of accessible information continues to expand in the information era, and big data has added the challenge of identifying relevant, appropriate, and useful information for strategy analysis.

Several game theoretic approaches such as the graph model for conflict resolution [3], argumentative analysis of options [4], and hierarchical graph models [5] have established themselves as useful frameworks for analyzing the strategic actions of multiple stakeholders in complex conflicts. However, few have focused on incorporating natural language into their analysis. On the other hand, in the field of natural language processing, research has enabled programs to identify and extract causal relations within sentences [6], although none have attempted the automated extraction of strategic relations.

Therefore, the objective of this research is to integrate natural language processing tools with conflict analysis methods to analyze the strategies of players involved in the creation of the AIIB. The research aims to automatically extract strategic expressions from a large corpus and convert the information into a visible model that can be analyzed with the graph model for conflict resolution [3], enabling the strategic analyses of large data sets. This paper presents the first steps to conducting automated strategic relation extraction from a corpus through the development of a natural language processing tool. The analytical implications of automated extraction are then considered by applying graph models of strategy relations created from the extracted sentences.

II. EXTRACTION OF STRATEGIC RELATIONS

Firstly, the method for extracting relations of strategies using natural language processing will be explained. The sentences containing clue expressions, or expressions that often represents strategic movements, were extracted from a corpus. The corpus used in this research is 404 articles published in the Mainichi Shimbun from July 4, 2014 to March 15, 2016 that included the word "AIIB". Machine learning is then used to determine whether or not the extracted sentence contains strategic relations.

A. Clue Expressions

Expression that works as a clue to find strategic relations from documents, or clue expression is used to extract strategic relations from Japanese news articles. Previously, Sakaji and Masuyama used clue expressions to extract causal relations from newspaper articles [6]. For example, in the sentence "If China establishes the AIIB, then America speeds up TPP negotiations", the expression "*If …, then*" is the clue expression. Six clue expressions (see Table 1) have been utilized to indicate the strategies contained within a sentence like "If player *A* does action *X*, then player *B* does action *Y*." This sentence structure forms a strategic relation, which refers to the relation of the first action (*A* does *X*) and reaction (*B* does *Y*).

In order to extract strategic relations in this manner, sentences that contain a) clue expressions, b) action *X*, and c) reaction *Y* must be identified. The six clue expressions were selected as appropriate because grammatical Japanese sentences would then place action *X* before and reaction *Y* after the clue expression. In order to identify these components within a sentence, each sentence must be broken down. Morphological analysis and dependency parsing are thus applied to the sentences containing clue expressions. Morphological analysis has been conducted with the Japanese part-of-speech and morphological analyzer MeCab[1], and dependency parsing conducted with a Japanese dependency structure analyzer CaboCha [7]. This resulted in 154 sentences being extracted from the corpus of 404 articles.

An issue that needed to be addressed was that clue expressions do not always co-occur with strategic expressions, e.g., "If China establishes AIIB, then AIIB will be one of the leading multilateral banks in Asia." Machine learning can be applied to address this issue to further automate the process of accurately identifying strategic relations within the corpus.

B. Supervised Machine Learning

To determine if sentences containing clue expressions actually contained causal relations, three syntactic features unique to the Japanese language and previously identified and utilized by Sakaji and Masuyama [6] have been extracted. To implement machine learning, the support vector classifier machine Libsvm [8] is used to implement supervised learning with previously prepared training data. The corpus used to compile the training data consists of 585 newspaper articles published by the Nikkei from July 4, 2014 to January 27, 2016 that included the word "AIIB". From this corpus, 188 sentences that contained clue expressions were extracted as training data. To determine whether the sentences in the training data had strategic relations in them, the following rules were set:

1. The sentence must contain the relation of *A*, *X*, *B*, *Y* of "If player *A* does action *X*, then player *B* does action *Y*."
2. Even if any of the *A*, *X*, *B*, *Y* is not clearly written in the sentence, if the meaning can be understood from the surrounding words, the sentence is valid
3. If either *A*, *X* or *B*, *Y* is expressed in demonstratives, the sentence is invalid.

Following the above rules, 64 of the 188 sentences in the training data were determined to contain strategic relations. After giving each of the sentences in the training data a tag of positive (to contain strategic relations) or negative (to not contain strategic relations), they are fed into the support vector machine with additional testing data to allow the machine to tag additional examples with positive or negative tags. The test data consisted of 154 sentences that contained clue expressions taken from the corpus.

C. Extraction of Strategic Relations and Evaluation

To evaluate the accuracy of the approach, the test data was manually examined to identify strategic relations in consistency with the rules mentioned above. The accuracy and coverage of this approach has been calculated as follows:

$$\text{Accuracy} = \frac{H \cap S}{H} \qquad (1)$$

$$\text{Coverage} = \frac{H \cap S}{S} \qquad (2)$$

H : Number of sentences extracted using machine learning

S : Number of sentences containing strategic relation within the sentences containing clue expressions

As summarized in Table 2, the support vector machine attained an accuracy rate of 60%. Analysis has shown that the features that have been incorporated into the machine learning process are insufficient in accurately identifying many strategic clue-expressions. The current research only applies syntactic features but future work will attempt to incorporate semantic features to improve accuracy.

III. Graph Model for Conflict Resolution and Strategic Maps

A. Application to Strategic Maps

After the sentences containing strategic relations were

TABLE I.
LIST OF CLUE EXPRESSIONS

Clue Expressions (in Japanese)	Meaning
Wo-ukete	In response to
Wo-uke,	In response to
sureba	If ~ does …
sitara	If ~ does …
Ni-taikou	Against ~
Heno-taikou	Against ~

TABLE III.
TEST DATA

Number of sentences that include clue expressions	154
Number of sentences containing strategic relations	45

TABLE II.
ACCURACY AND COVERAGE

Number of sentences that actually includes relations of strategies among the sentences extracted (H ∩ S)	30
Number of sentences extracted (S)	50
Number of sentences containing strategic relations(H)	45
Accuracy	0.600
Coverage	0.667

[1] http://taku910.github.io/mecab/

extracted, the sentences were broken down into their action and reaction. Then, each action's player and the content of the action were extracted using dependency parsing. This player-action pair is defined as a strategy. Only three sentences were able to identify player-action pairs accurately. One example of a successful extraction is of the sentence "If Taiwan strengthens its relation with China, then China will increase its claim on the South China Sea." The player of the first action is Taiwan, the action to strengthen its relation with China. The player of the reaction is China and the action is to increase claim on the South China Sea.

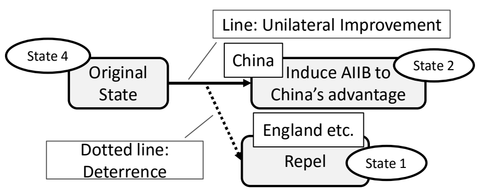
Using the extracted strategies (player-action pairs), a strategic map can be drawn and used for analysis. Horita (2000) introduced the strategic map to represent social conflicts under a high degree of disagreement [4]. Strategic maps are drawn by placing strategy in a box and drawing arrows to show the relation between strategies.

B. *Application to the Graph model for Conflict Resolution*

The graph model for conflict resolution (GMCR) [6] enables analysis of strategic movements and counter-movements in a conflict to predict the most likely outcome. The information needed to construct a graph model are (adapted from [6]):

1. Decision Makers (DM) in the conflict
2. Options under the control of each DM
3. Relative preferences for each DM
4. Infeasible states
5. Allowable transitions among states

A graph model can thus be constructed using the information contained within the strategic maps. That is, each state in a GMCR corresponds to a strategy in the strategic map. The arrows stand for improvements, which are a transition of states that lead to more preferable outcomes provided that all of the other players' choices do not alter. The broken line in the figure represents sanctions to deter against possible unilateral improvements for a player.



DM	Option				
China	Induce AIIB to China's advantage	Y	Y	N	N
England etc.	Repel	Y	N	Y	N
State		1	2	3	4

Figure 1. Example of strategic map and the set of corresponding states

An example is shown in figure 1. This figure shows a strategic map and a table of possible states. There are two DMs, China and England. China has the option to use AIIB to maximize China's interest. England has the option to repel against China's action. Y stands for undertaking the option, whereas N stands for not taking the option, resulting in 4 possible states. If China prefers state 4 to state 1, and England prefers state 4 to state 2, the Nash equilibrium is state 4. Some of the preference information can be retrieved from logically interpreting the extracted texts. E.g., "If *A* does *X*, then *B* does *Y*", implies that that player *B* prefers *Y* to *X*.

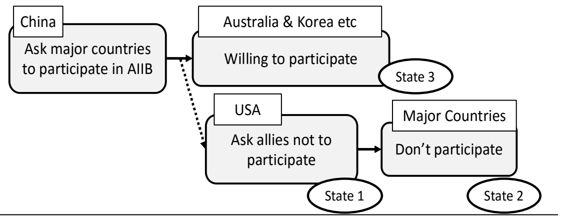
Another analysis method derived from GMCR is inverse GMCR [9]. To conduct an inverse GMCR analysis, the desired resolution for each player is necessary, and information on the relative preferences of each DM is not required. Inverse GMCR can deal with cases where the preference information is not certain. Using the same example of figure 1, if the desired state is state 1 for England, it can be derived that the preference ranking should be (state 4) $\succ$ (state 1) for China. When interpreting text data, if it is written that a strategy has been acted upon, that strategy could be considered a Nash Equilibrium, and thus preference information can be extracted from this assumption. Further use of these logical interpretations and assumptions being applied to the extracted texts would allow researchers to automate the process of preference extraction and strategic analysis.

Applying the strategic map to the GMCR allows us to interpret the actions of China and England, namely, that China would try to manage AIIB in order to maximize its own interest, but in the face of a deterrence threat from England, China would not take the first choice to induce the AIIB.

IV. Strategy Analysis Using Strategic Maps

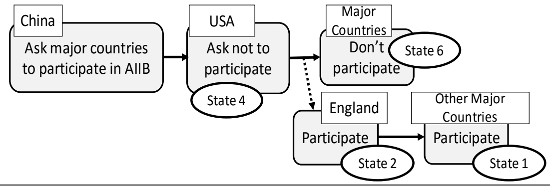
In this manner a summative strategic map for the 585 newspaper articles used in section II as training data was constructed, and GMCR was applied to interpret several sub-games within the map. This section will provide some examples of strategic analyses that were made possible by integrating the two approaches. Figure 2 shows the situation before England decided to join the AIIB, whereas figure 3 represents the situation after England announced its participation. Figure 2 shows that before England joined the AIIB, news articles claimed that other major countries did not accept invitations to join the AIIB as a result of US pressure. However, figure 3 describes how after England announced its participation, other countries changed their decisions and also joined the AIIB. The model implies that the major countries' preferences have been changed by England's intervention.

While this is just one example, by constructing a strategic map using a large corpus, the complex relations between events that at first do not seem inter-related become clear. In this case, as shown in figure 4, connections between the AIIB, TPP negotiations, territorial disputes in the South China Sea, and the annexation of Crimea by the Russian Federation were revealed. Complex geopolitical strategies and analyses were elucidated, such as claims that China's increase in territorial claim of the South China Sea are a result of US preoccupation with Russia's annexation of Crimea and terrorism. The territorial disputes are listed alongside the



DMs	Option				
USA	Prevent participation	Y	Y	N	N
Major countries	Participate	Y	N	Y	N
	State	1	2	3	4

Figure 2. Strategic map and states before England joined AIIB



DMs	Options								
England	Participate	Y	Y	Y	Y	N	N	N	N
USA	Prevent	Y	Y	N	N	Y	Y	N	N
Other Major Countries	Participate	Y	N	Y	N	Y	N	Y	N
	State	1	2	3	4	5	6	7	8

Figure 3. Strategic map and states after England joined AIIB

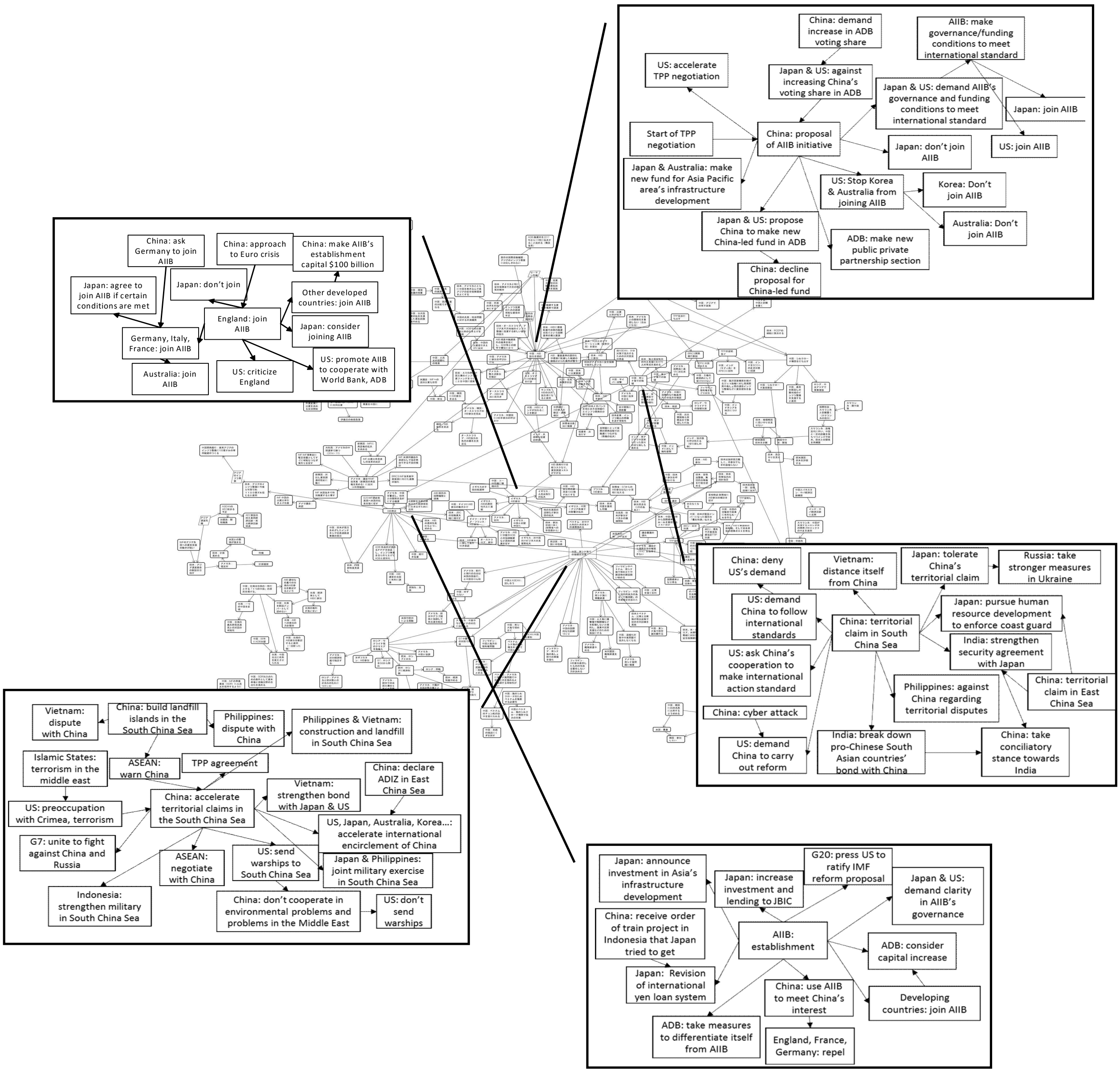


Figure 4. Example of strategic map using large corpus

AIIB initiative as the main causes of the US's acceleration of the TPP negotiations.

Challenges remain in automating the creation of the strategic map. Few parts of the strategic map were accurately reflected in strategies extracted using the natural language processing methods explained in section II (e.g., Figure 5).

V. CONCLUSION

This paper has presented the first step at attempts at automated strategic relation extraction through the development of a natural language processing tool for analyzing large data sets. The analytical implications of such a tool has been considered by applying the extracted information to Graph models. A refined tool would enable the strategic analysis of large amounts of text data concerning complex social and political conflicts such as issues surrounding the AIIB. Two tasks that attempt to achieve these objectives has been presented:

1. Extraction of strategic relations from a corpus using natural language processing
2. Application of extracted strategies to a strategic map and strategic analysis of the map using GMCR

Further integration of these two approaches will lead to more accurate automated extraction of strategic relations within large data sets, as well as the eventual automation of strategic analyzing large data sets.

ACKNOWLEDGMENTS

The authors would like to thank K. Ozawa of the University of Tokyo for providing valuable advice in conducting this research. They would also like to thank H. Suhara of VeriServe Corporation and K. Inui of Tohoku University for taking the time to give insights about Japanese natural language processing.

REFERENCES

[1] M. Hoshino, "Asian infrastructure investment bank (AIIB) and TPP," *Bulletin of Niigata Sangyo University Faculty of Economics*, vol. 45, pp. 25–40, June 2015.

[2] The White House – President Barack Obama. (2005, April 17), Remarks by President Obama and Prime Minister Renzi of Italy in Joint Press Conference. Retrieved from https://www.whitehouse.gov/

[3] L. Fang, K. Hipel, and M. Kilgour, *Interactive Decision Making: the Graph Model for Conflict Resolution*. Wiley-Interscience, September 1993

[4] M. Horita, "Folding arguments: a method for representing conflicting views of a conflict", *Group Decision and Negotiation,* vol.9, 2000, pp.63-83

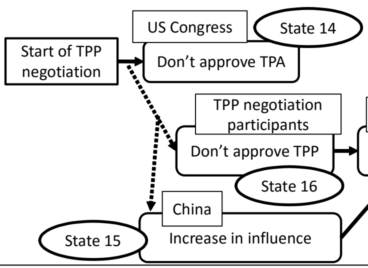
[5] S. He, M. Kilgour, and A. Bashar, "A basic hierarchical graph model for conflict resolution with application to water diversion conflicts in China," *INFOR Journal*, vol.51, no.3, pp.103-119, August 2013.

[6] H. Sakaji and S. Masuyama, "A method for extracting sentences including causal relations from newspaper articles," *The IEICE transactions on information and systems (Japanese edition)*, J94-D(8), pp. 1496–1506, August 2011.

[7] T. Kudo and Y. Matsumoto, "Japanese dependency analysis using cascaded chunking," *CoNLL 2002: Proceedings of the 6th Conference on Natural Language Learning 2002 (COLING 2002 Post-Conference Workshops)*, 2002, pp.63–69

[8] C. Chang and C. Lin, "Libsvm : a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology* , vol.2, 2011, pp. 27:1-27:27

[9] R. Kinsara, M. Kilgour, and K. Hipel, "Inverse approach to the graph model for conflict resolution," *IEEE Transactions on Systems, Man, and Cybernetics-Systems,* vol.45, no.5, pp.734-742, May 2015



Player	Option																
US President	Sign TPP	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N	N	N	N
US Congress	Approve TPA	Y	Y	Y	Y	N	N	N	N	Y	Y	Y	Y	N	N	N	N
TPP negotiation participants	Approve TPP	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N
China	Increase influence	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N
	State	1	**2**	3	4	5	6	7	8	9	10	11	**12**	13	**14**	**15**	**16**

Figure 5. Strategic Map and set of States of strategies automatically extracted