An Extraction Method for Future Reference Expressions Using Morphological and Semantic Patterns

Yoko NAKAJIMA[1] Michal PTASZYNSKI[2]
Hirotoshi HONMA[1] Fumito MASUI[2]

Abstract: When we categorize textual data according to time categories the three main types of information that come up are the past, the present, and the future. In this paper we present our study in predicting the future. In particular, we aim at detecting expressions which refer to future events (temporal expressions, etc.) and apply them to support the prediction of probable future outcomes. In order to realize the future prediction support, we firstly need to be able to find out whether a sentence refers to the future or not in general. We propose a method of bi-polar text classification for sentences into either future-related or non-future-related (other). To do this we use a machine learning based sentence pattern extraction system SPEC and report on the accuracy of extracted patterns. We train the classifier using morphological and semantic representations of sentences and show that it is possible to extract fully automatically frequent patterns from sentences referring to the future.
Key words: Information Extraction, NLP, Future Prediction, Semantic Role;

1 Introduction

In everyday life people use past events and their own knowledge to predict future events. To obtain the necessary data for such everyday predictions, people use widely available sources of information (newspapers, Internet). In my study I focus on sentences that make reference to the future. Below is an example of a future-reference sentence published in a newspaper[1] (translation by the author),

- *Science and Technology Agency, the Ministry of International Trade and Industry, and Agency of Natural Resources and Energy conferred on the necessity of a new system, and decided to set up a new council.*

The sentence claims that the country will construct a new energy system. Interestingly, despite the sentence is written with the use of past tense ("conferred", "decided") the sentence itself refers to future events ("setting up a new council"). Such references to the future contain information (expressions, patterns, causal relations) relating it to the specific event that may happen in the future. The prediction of the event depends on the ability to recognize this information.

A number of studies have been conducted on the prediction of future events with the use of time expressions [2, 6], SVM (bag-of-words) [1], causal reasoning with ontologies [10], or keyword-based linguistic cues ("will", "is going to", etc.) [5]. In my research I assumed that the future reference in sentences occurs not only on the level of surface (time expressions, words) or grammar, but consist of a variety of patterns both morphological and semantic.

[1]Department of Information Engineering, National Institute of Technology, Kushiro College
[2]Department of Computer Science, Kitami Institute of Technology
[1]Japanese daily newspaper *Hokkaido Shinbun*.

The outline of this paper is as follows. Section 2 presents our investigation into future reference expressions. Section 3 describes the proposed method. Section 4 describes the experiments evaluating the method. In section 5 we verify the performance of the method on new validation set and compare it to the state-of-the-art. Finally, we conclude the paper and present some of the possible directions for improvement and application of the developed method in section 6.

2 Investigation in Future Reference Expressions

We performed a study of expressions which refer to a change in time in general or to the future in particular. The study has been performed by reading through articles from the following newspapers: *the Nihon Keizai Shimbun*[2], *the Asahi Shimbun*[3], *the Hokkaido Shimbun*[4]. We used all newspapers in their both paper and Web version. From the above newspapers we manually extracted from various articles 270 representative sentences which referred to the future. Next, on the sentences we manually annotated future expressions. There were 70 time-related expressions and 141 unique future expressions (words, phrases, etc.) that were not time-related.

Some examples of the expressions are represented in Table 1. There are two kinds of future-related expressions. First consists of concrete expressions which include numerical values, such as "year 2013", or "11 o'clock".

[2]http://www.nikkei.com/
[3]http://www.asahi.com/
[4]http://www.hokkaido-np.co.jp/

Table 1: Examples of future- and time-related expressions.

Type of expression	Number found	Examples; Y=year, M=month (usually appearing as numerical values)
Time-related expressions	70	*Y-Nen M-gatsu kara* ("from month M year Y"), *kongo Y-nenkan ni* ("in next Y years"), *Y-gatsu gejun ni mo* ("late in year Y"), etc.
Future expressions	141	*mezasu* ("aim to") (11), *hōshin* ("plan to") (12), *mitooshi* ("be certain to") (9), *kentō* ("consider to") (9), *-suru* ("do") (76), *-iru* ("is/to be") (36), etc.

Table 2: An example of semantic representation of words performed by ASA.

Surface	Semantic (Semantic role, Category, etc.) and grammatical representation
mezasu ("aim to")	No change (activity)-action aiming to solve [a problem]-pursuit; Verb;
hōshin ("plan to")	Other;Noun;
mitooshi ("be certain to")	Action;Noun;
kentō ("consider to")	No change (activity)-action aiming to solve [a problem]-act of thinking;Noun;
-suru ("do")	Change-creation or destruction-creation (physical);Verb;
-iru ("is/to be")	Verb;

Second is derived from grammatical information (verb tense, word order, particles, etc.), such as phrases "will [do something]", "the middle of a month", "in the near future", or particles *-ni* ("in, due, till", point of time), *-made* ("until", implied deadline for continuous action), or *-madeni* ("until", implied deadline for single action).

However, many of the extracted 270 sentences did not contain typical time or future expressions. Among all expressions we annotated on the sentences, 55% appeared two or more times, while 45% only once. This is due to the fact that there could be many variations of expressions, which could function as future expressions only in a specific context. However, we can assume that these which appear the most often have a characteristics of future expressions. Therefore if we consider sentences and their different representations (grammatic, semantic) as sets of patterns which occur in a corpus (collection of sentences/documents) we should be able to extract from those sentences new patterns referring to the future. For example, a sentence annotated with semantic roles should provide semantic patterns occurring frequently in future-reference sentences. Below we describe the method to extract such patterns.

3 Future Reference Pattern Extraction Method

3.1 Morphosemantic Patterns

In the first stage, all sentences included in the datasets (see section 4.1), are represented in **morphosemantic patterns** (MoPs).

The idea of MoPs has been described widely in linguistics and structural linguistics. For example, [8] distinguish

Table 3: An example of a sentence analyzed by ASA.

Example I: Romanized Japanese (RJ): *Ashita kare wa kanojo ni tegami o okuru darō.* / **Glosses:** Tomorrow he TOP her DIR letter OBJ send will (TOP: topic particle, DIR: directional particle, OBJ: object particle.) / **English translation (E):** He will [most probably] send her a letter tomorrow.

No.	Surface	Label
1	*ashita*	[Time-Point]
2	*kare ha*	[Agent]
3	*kanojo ni*	[Patient]
4	*tegami o*	[Object]
5	*okuru darou*	[State_change]-[Place_change]-[Change_of_place(physical)]

them as one of the two basic types of morphological operations on words, which modify the Lexical Conceptual Structure (LCS), or the semantic representation of a word. As for practical application of the idea, [7] applied MoPs to analyze an Indonesian suffix *–kan*. Later [3] applied MoPs to improve links between the synsets in WordNet. More recently [11] used MoPs to analyze a lexicon in Croatian, a language rich both morphologically and semantically. In this research we used datasets in Japanese, and applied MoPs for the same reason. Using only one representation narrows the spectrum of analyzed information. Moreover, till now there has been no practical application of MoPs to solving real-world problems. In this paper we present the first attempt of this kind.

We generated the morphosemantic model using semantic role labeling with additional morphological information. Below we describe in detail the process of morphosemantic representation of sentences.

At first, the sentences from the datasets are analyzed using semantic role labeling (SRL). SRL provides labels for words and phrases according to their role in sentence context. For example, in a sentence "John killed Mary" the labels for words are as follows: John=`actor`, kill[past]=`action`, Mary=`patient`. Thus the semantic representation of the sentence is "`actor-action-patient`".

For semantic role labeling in Japanese we used **ASA**[5], a system, developed by [12], which provides semantic roles for words and generalizes their semantic representation using an originally developed thesaurus. Examples of labels ASA provides for certain words are represented in Table 2. An example of SRL provided by ASA is represented in Table 3.

Moreover, not all words are semantically labeled by ASA. The omitted words include those not present in the thesaurus, as well as grammatical particles, or function words not having a direct influence on the semantic structure of the sentence, but in practice contributing to the overall meaning. For such cases we used a morphological analyzer MeCab[6] in combination with ASA to provide morphological information, such as "Proper Noun", or "Verb". However, in its basic form MeCab provides morphological information for all words separately. Therefore, there often occurs a situation where a compound

[5] http://cl.it.okayama-u.ac.jp/study/project/asa

[6] http://code.google.com/p/mecab/

word is divided. For example "Japan health policy" is one morphosemantic concept, but in grammatical representation it takes form of "Noun Noun Noun". Therefore as a post-processing procedure we added a set of linguistic rules for specifying compound words in cases where only morphological information is provided.

Moreover, as it is shown on Table 3, some labels provided by ASA are too specific. Therefore in order to normalize and simplify the patterns, we specified the priority of label groups in the following way.

1. Semantic role (Agent, Patient, Object, etc.)
2. Semantic meaning (State_change, etc.)
3. Category (Dog $\rightarrow$ Living animal $\rightarrow$ Animated object)
4. In case of no analysis by ASA perform compound word clustering for parts of speech (e.g., "International Joint Conference on Artificial Intelligence" $\rightarrow$ `Adjective Adjective Noun Preposition Adjective Noun` $\rightarrow$ `Proper_Noun`)

Furthermore, post-processing in the case of no semantic information is organized as follows.

- If a compound word can be specified, output the part-of-speech cluster (point 4 above).
- If it is not a compound word, output part-of-speech for each word.

Below is an example of a sentence generalized with the semantic role labeling method applied in this research.

Romanized Japanese: *Nihon unagi ga zetsumetsu kigushu ni shitei sare, kanzen yōshoku ni yoru unagi no ryōsan ni kitai ga takamatte iru.*

English: As Japanese eel has been specified as an endangered species, the expectations grow towards mass production of eel in full aquaculture.

SRL: `[Object][Agent][State_change][Action][Noun][State_change][Object][State_change]`

3.2 Automatic Extraction of Frequent Patterns

Having all sentences represented in morphosemantic structure, we used SPEC, a system for extraction of sentence patterns [9]. SPEC is a system automatically extracting frequent sentence patterns distinguishable for a corpus (a collection of sentences). Firstly, the system generates ordered non-repeated combinations from all sentence elements. In every n-element sentence there is k-number of combination groups, such as that $1 \leq k \leq n$, where k represents all k-element combinations being a subset of n. The number of combinations generated for one k-element group of combinations is equal to binomial coefficient, represented in equation 1. In this procedure the system creates all combinations for all values of k from the range of $\{1, ..., n\}$. Therefore the number of all combinations is equal to the sum of all combinations from all k-element groups of combinations, like in the equation 2.

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (1)$$

$$\sum_{k=1}^{n} \binom{n}{k} = \frac{n!}{1!(n-1)!} + \frac{n!}{2!(n-2)!} + ... + \frac{n!}{n!(n-n)!} = 2^n - 1 \quad (2)$$

Next, the system specifies whether the elements appear next to each other or are separated by a distance by placing a wildcard ("*", asterisk) between all non-subsequent elements. SPEC uses all patterns generated this way to extract frequent patterns appearing in a given corpus and calculates their weight. The weight can be calculated in several ways. Two features are important in weight calculation. A pattern is the more representative for a corpus when, firstly, the longer the pattern is (length k), and the more often it appears in the corpus (occurrence O). Thus the weight can be calculated by

- awarding length (`LA`),
- awarding length and occurrence (`LOA`),
- awarding none (normalized weight, `NW`).

The normalized weight w_j is calculated according to equation 3. Normalization is performed to make weights fit in range from +1 to -1, and is achieved by subtracting 0.5 from the initial score and multiplying the intermediate product by 2.

$$w_j = \left(\frac{O_{pos}}{O_{pos} + O_{neg}} - 0.5\right) * 2 \quad (3)$$

The generated list of frequent patterns can be also further modified. When two collections of sentences of opposite features (such as "future-related vs. non-future-related") is compared, the list will contain patterns that appear uniquely in only one of the sides (e.g., uniquely positive patterns and uniquely negative patterns) or in both (ambiguous patterns). Thus pattern list can be modified by

- using all patterns (`ALL`),
- erasing all ambiguous patterns (`AMB`),
- erasing only those ambiguous patterns which appear in the same number in both sides (zero patterns, `0P`).

Moreover, a list of patterns will contain both the sophisticated patterns (with disjoint elements) as well as more common n-grams. Therefore the system can be trained on a model using

- patterns (`PAT`), or
- only n-grams (`NGR`).

All combinations of those modification are tested in the experiment.

4 Evaluation Experiment

4.1 Dataset Preparation

From all collected sentences referring to future events (section 2) we randomly selected 130 sentences and manually collected another 130 sentences which did not make any reference to the future (describing past, or present events). Out of those sentences we created two experiment sets. The first one containing 100 sentences, with 50 future-reference sentences and 50 non-future-reference sentences (later called "set50"). The second one containing 260 sentences, also with equal distribution of sentences of the two types (later called "set130"). All sentences were represented in morphosemantic structure according to the procedure described in section 3.1. From the sentences preprocessed this way we extracted pattern lists using the extraction procedure described in section 3.2.

4.2 Experiment Setup

We designed the experiment as a text classification task with the prepared datasets applied into 10-fold cross validation. The classification was performed as follows. Each test sentence was given a score calculated as a sum of weights of patterns extracted from training data and found in the input sentence (equation 4).

$$score = \sum w_j, (1 \geq w_j \geq -1) \quad (4)$$

The results were calculated using standard Precision, Recall and balanced F-score. However, if the initial collection of sentences was biased toward one of the sides (e.g., sentences of one kind are in larger number or longer), there will be more patterns of a certain type. Thus, using a rule of thumb in evaluation (e.g., fixed threshold above which a sentence is classified as either future-related or not) does not provide sufficiently objective view on results. Therefore we additionally performed threshold optimization to find which modification of the classifier achieved the highest scores. In the experiment 14 different versions of the classifier are compared under 10-fold cross validation condition. Since the experiment was performed on two datasets, we obtained overall 280 experiment runs. There were several evaluation criteria. Firstly, we looked at top scores within the threshold span. Secondly, we checked which version got the highest break-even point (BEP) of Precision and Recall. Finally, we checked the statistical significance of the results using paired t-test.

4.3 Classification Results

Experiment results (F-score) for all classifier versions tested on set50 and set130 for models trained on n-grams and patterns are compared separately in Figures 1, 2, 3, and 4.

For most cases pattern-based approach obtained significantly higher scores than n-grams, which means that there are meaningful frequent patterns in future sentences, more sophisticated than n-grams. When it comes to the modifications of pattern lists and weight calculations, only deleting zero patterns did not significantly influence the results. A larger difference was observable when all ambiguous

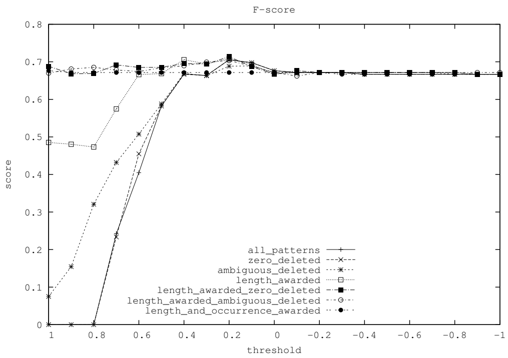

Figure 1: Results (F-score) for all classifier versions tested in the experiment on set50 for model trained on patterns.

patterns were deleted. Moreover, awarding pattern length in weight calculation always yielded better results. The highest results achieved were F=0.71 with P=0.56 and R=0.98 for the version of the classifier which used pattern list with zero-patterns deleted and length awarded. The greatest improvement of patterns in comparison with n-grams was always in Recall, which means that there are valuable patterns omitted in the model trained only on n-grams. Precision does not change significantly and oscillates around 0.55–0.60. This means that the point of around 0.55–0.60 is the optimal maximum that could be achieved with the morphosemantic patterns we used in this study. In the future we will look for an improvement improving Precision while not reducing Recall.

When it comes to the highest achieved scores, the highest F-score for patterns was 0.71, while for n-grams it was 0.70. Although the difference is not large, patterns, due to better Recall usually achieve high F-score for most of the threshold, where n-grams usually score lower (compare Figures 1 with 3, and Figure 2 with 4).

Next we compared in detail the results between the two datasets, **set50** and **set130**. For set50, the F-score reached plateau at around 0.67–0.71 for patterns and 0.67–0.70 for n-grams. For set130 the plateau for F-score was reached at around 0.67–0.70 for patterns and 0.67–0.69 for n-grams. The optimal threshold (from range 1.0 to -1.0) was around 0.0. which means both sides of the training set were balanced.

The F-scores for the version of the classifier using pattern list with all ambiguous patterns deleted performed better than other pattern list versions (unmodified and zero deleted), although the differences were not quite statistically significant ($p<0.06$). The performance was generally better when the length of patterns was used to modify weight calculation. Especially both modified versions of the classifier (without zero-patterns and without all ambiguous patterns) retained high F-score thorough all threshold. Applying pattern length in weight calculation yielded better results within the specified threshold. Also, the performance for the algorithm as a whole is similar for set50 and set130. Larger dataset usually contains more ambiguities, thus the results would be expected to de-

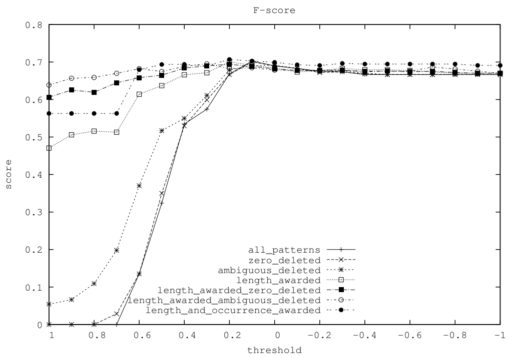

Figure 2: Results (F-score) for all classifier versions tested in the experiment on set130 for model trained on patterns.

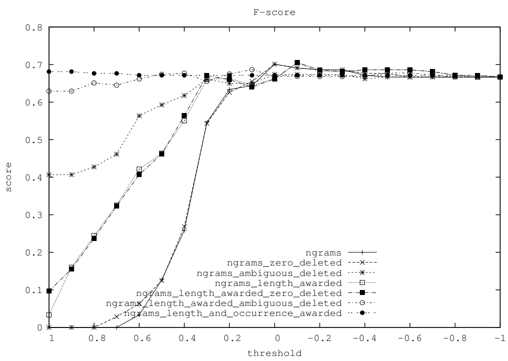

Figure 3: Results (F-score) for all classifier versions tested in the experiment on set50 for model trained on n-grams.

grade. With the proposed approach the differences are negligible and statistically not significant.

5 Method Validation

5.1 Performance Change for Small Pattern Sets

At first we performed estimation of the effectiveness of morphosemantic patterns in future reference sentence classification. Firstly, we collected the following additional new validation set, unrelated to the initial datasets. From one year (1996) of *Mainichi Shinbun* newspaper we extracted 170 sentences from articles appearing on first three pages of each edition, and articles from the topics "economy", "international events" and "energy."

We manually annotated these sentences as either future or non-future related with five annotators: one expert annotator and four laypeople. Each sentence was annotated by one expert- and two layperson-annotators. We decided to leave the sentences for which there was an agreement between at least one layperson annotator and the expert. In result 59% (exactly 100 sentences) were left as the validation set.

Next, we classified these newly obtained sentences using the most frequent patterns (first 5 of them are repre-

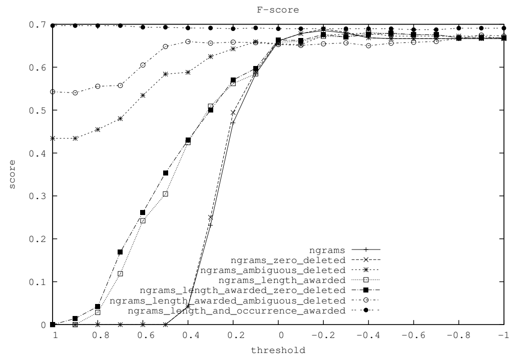

Figure 4: Results (F-score) for all classifier versions tested in the experiment on set130 for model trained on n-grams.

sented in Table 4) generated in previous experiment. In particular, we performed pattern matching on the new sentences with the following sets:

A: first 10 patterns,
B: adding 5 patterns longer than three elements to set A,
C: subtracting 5 patterns from the tail of set A (to discard less frequent patterns shorter than three elements),
D: using only first 10 patterns containing more than three elements (differently to Set A which contains also frequent but shorter patterns).

Once performance reached plateau (F-score = 0.43), increasing the number of patterns made little difference. The performance of pattern set C was poor since only a few patterns are used. The Precision of pattern set D is slightly higher than that of the other sets. This indicates it could be more effective to use frequent morphosemantic patterns containing more than three elements, even when the number of applied patterns is small. From the above, we conclude that it would be more effective to use patterns consisting of a few (two or three) elements if the focus of the extraction was on Recall, whereas it would be more effective to use patterns consisting of three or more elements if the focus was on Precision.

The scores in this experiment were lower than in the evaluation experiment. However, we were able to extract future reference sentences with approximately 40% of Precision using only ten patterns, a score not far below the one achieved in the evaluation experiment (in which a total of 1102 patterns was used). This suggests that the performance could be also further improved when morphosemantic patterns are narrowed to those appearing in specific genre of events (only "economy", or only "energy").

5.2 Compariosn with State-of-the-Art

We also compared our experimental results with those reported by [4]. In their experiment they extracted future reference sentences with 10 words and phrases unambiguously referring to the future, such as temporal expressions like "will," "may," "be likely to", etc. We translated those

Table 4: Examples of extracted morphosemantic patterns.

Occ.	Future Reference Patterns	Occ.	Non-future Reference Patterns
26	[Action]*[State change]	5	[Place]*[Agent]
43	[Action]*[Object]	4	[Numeric]*[Agent]
42	[Action]*[Action]	4	[Verb]*[Artifact]
20	[State change]*[Object]	4	[Person]*[Place]
16	[State change]*[State change]	3	[Numeric]*[Agent]*[Action]
	⋮		⋮

Table 5: Comparison of results for validation set between different pattern groups and the state-of-the-art.

Pattern set	Precision	Recall	F-score
10 patterns	0.39	**0.49**	**0.43**
15 patterns	0.38	**0.49**	**0.43**
5 patterns	0.35	0.35	0.35
10 pattern with only over 3 elements	**0.42**	0.37	0.40
Optimized (see Figure 5)	**0.76**	**0.76**	**0.76**
[4]	**0.50**	0.05	0.10

phrases into Japanese and applied to the new validation dataset of 170 sentences. The results of were low with P = 0.50, R = 0.05, F = 0.10. Although the Precision seems higher than the one described in section 5.1, our method extracted correctly much more future referring sentences with only 10 morphosemantic patterns. This indicates that the proposed method is valid. The reason for the low score obtained by the method of [4] on our validation dataset, despite its showing better performance previously could be explained by the differences in the approach. [4] used future-related patterns well known in linguistics, and searched for future sentences on the Internet which contains sufficient amount of data for extraction with even minimal number of seed words. We on the other hand trained our method automatically without providing any linguistic knowledge on a corpus from which we automatically extracted sophisticated morphosemantic patterns.

5.3 Performance of Fullly Optimized Model

Finally, we verified the performance of the fully optimized model. The results of evaluation experiment (section 4) indicated that the model with the highest overall performance was the one using pattern list containing all patterns (including both ambiguous-, zero-patterns and n-grams) with weights modified by awarding pattern length. We retrained the above model using all sentences from set130 and verified the performance by classifying the new validation set of 100 sentences.

As the evaluation metrics we used standard Precision, Recall and F-score. The scores of sentences oscillated from -0.01 to 2.27. The stronger was morphosemantic similarity to the training data the higher was the score. We also verified the performance for each threshold, beginning form 0.0 and checked every 0.2, up-till 2.2. The overall performance is represented in Figure 5. The highest reached Precision was 0.89, at R=0.13 with F=0.22.

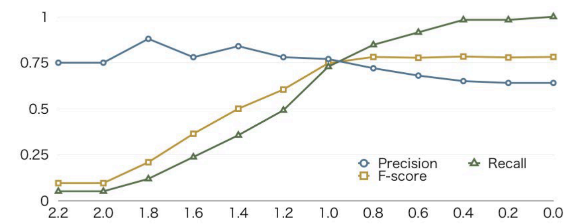

Figure 5: The results (F-score, Precision and Recall) for classification of future reference sentences in the test data.

The highest reached F-score was 0.78 with Precision = 0.65 and Recall = 0.98 around the threshold of 0.4. Finally, break-even point (BEP) was at 0.76, which indicates that the proposed method trained on automatically extracted morphosemantic future reference patterns is sufficiently capable to classify future reference sentences.

Apart from the automatic classification results, we were also interested in the actual patterns that influenced the results. In Figure 6 we present detailed analysis of two sentences which obtained high scores in the experiment with first four patterns mapped on the sentences to facilitate better understanding of the future-referring morphosemantic patterns.

6 Conclusions and Future Work

We presented a novel method for extracting references to future events from news articles, based on automatically extracted morphosemantic patterns. The method firstly represents news articles in morphosemantic structure using semantic role labeling supported with part-of-speech tagging. Next, it extracts all possible morphosemantic patterns from the corpus including sophisticated patterns with disjoint elements. After being trained on both future- or non-future-related patterns we performed a text classification experiment in which

we compared 14 different classifier versions to chose the optimal settings. The optimized method was further validated on completely new dataset, and compared to the state-of-the-art. The proposed method outperformed the state-of-the-art and when optimized reached the final score of high Precision and Recall with break even point and plateau balanced on 76%.

In the future we plan to increase the size of the experimental datasets to evaluate the method more thoroughly and determine a general morphosemantic model of future reference sentences. This would be useful in estimating probable unfolding of events, and would contribute to the task of trend prediction in general.

1. Score=2.27

RJ *Dōsha wa kore made, Shigen Enerugī-Chōni taishi, dō hatsudensho no heisa, kaitai ni tsuite hōshin o setsumei shite kitaga, kaitai ni tsuite no hōteki kisei wanai tame, dōchō mo kaitai no kettei o shitatameru koto ni nari-sōda.*

E So far the company has been describing to the Agency for Natural Resources and Energy the policy for either closure or dismantling of the plant, and since there are no legal regulations found for dismantling, it is most likely that the agency will also lean to the decision of dismantling.

MS `[Agent][Other][Organization][Action][State-change][State-change][Object][Role][State-change][State-change][Action][Adjective][Thing][Agent][State-change][Other][Verb]`

MoPs `[Agent]*[Verb]`,
`[Agent]*[Organization]*[Verb]`,
`[Agent]*[Action][State-change]*[Verb]`,
`[Agent]*[Organization]*[State-change]*[Verb]`.

2. Score=1.77

RJ *Ippō, senkyo kikan-chū ni '400 man-ri wokoeru Jimintō shiji no shomei o atsume, ōen shita'(tō kanbu) to iwa reru sekiyu, gasu nado enerugī kanren dantai ni taisuru hiaringu de wa, kūki ga ippen.*

E On the other hand, saying during the elections that they "collected the signatures of more than 4 million people supporting the Liberal-Democratic Party" (citation after the party's leader), during the hearing for the organizations related to the energy [sources] such as oil and gas, completely changed the atmosphere.

MS `[Action][Action][Numeric][Verb][Action][Object][State-change][No-State-change-activity][Citation][Verb][Thing][Thing][Action][Action][Object][State-change]`

MoPs `[Action]*[State-change]`,
`[Action]*[State-change]*[Object][State-change]`,
`[Action]*[State-change]*[State-change]`,
`[Action]*[Action]*[State-change]*[State-change]`.

Figure 6: Examples of two sentences which obtained high scores in the experiment with their morphosemantic structure and extracted morphosemantic patterns. Each example contains in order: Score, Romanized Japanese [RJ], English Translation [E], Morphosemantic structure [MS], Morphosemantic future-reference patterns found in this sentence [MoPs]; for each example sentence, three examples of patterns from the list they contain (MoPs) are underlined, double underlined, overlined, or highlighted in gray.

References

[1] Eiji Aramaki, Sachiko Maskawa, Mizuki Morita. 2011. Twitter catches the flu: Detecting influenza epidemics using twitter. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 1568–1576.

[2] R. Baeza-Yates. 2005. Searching the Future. *SIGIR Workshop on MF/IR*.

[3] Christiane Fellbaum, Anne Osherson, and Peter E. Clark. 2009. Putting semantics into WordNet's "morphosemantic" links. *Human Language Technology. Challenges of the Information Society*. Springer Berlin Heidelberg. pp. 350-358.

[4] Adam Jatowt, Ching-man Au Yeung. 2011. Extracting collective expectations about the future from large text collections. *Proceedings of the 20th ACM international conference on Information and knowledge management*, pp. 1259–1264.

[5] Adom Jatowt, Hideki Kawai, Kensuke Kanazawa, Katsumi Tanaka, Kazuo Kunieda, Keiji Yamada. 2013. Multi–lingual, Longitudinal Analysis of Future-related Information on the Web. *Proceedings of the 4th International conference on Culture and Computing 2013*, IEEE Press.

[6] Kensuke Kanazawa, Adam Jatowt, Satoshi Oyama, Katsumi Tanaka. 2010. Exracting Explicit and Implicit future-related information from the Web(O) (in Japenese). *DEIM Forum 2010*, paper ID: A9-1.

[7] Paul Kroeger. 2007. Morphosyntactic vs. morphosemantic functions of Indonesian *-kan*. In *Architectures, Rules, and Preferences: Variations on Themes of Joan Bresnan*, edited by A. Zaenen, J. Simpson, T. H. King, J. Grimshaw, J. Maling and C. Manning, pp. 229–251. Stanford, CA: CSLI Publications.

[8] Beth Levin and Malka Rappaport Hovav. 1998. *Morphology and Lexical Semantics*. In Spencer and Zwicky, eds., pp. 248–271.

[9] Michal Ptaszynski, Rafal Rzepka, Kenji Araki and Yoshio Momouchi. 2011. Language combinatorics: A sentence pattern extraction architecture based on combinatorial explosion. *International Journal of Computational Linguistics (IJCL)*, Vol. 2, Issue 1, pp. 24-36.

[10] Kira Radinsky, Sagie Davidovich and Shaul Markovitch. Learning causality for news events prediction. *Proceedings of the 21st International Conference on World Wide Web*. ACM, 2012.

[11] Ida Raffaelli. 2013. The model of morphosemantic patterns in the description of lexical architecture. *Lingue e linguaggio*, Vol. 12, No. 1 (2013), pp. 47-72.

[12] Koichi Takeuchi, Suguru Tsuchiyama, Masato Moriya, Yuuki Moriyasu. 2010. Construction of Argument Structure Analyzer Toward Searching Same Situations and Actions. *IEICE Technical Report*, Vol. 109, No. 390, pp. 1-6.