A Class-oriented Approach to Building a Paraphrase Corpus

Atsushi Fujita
Graduate School of Informatics,
Kyoto University
fujita@pine.kuee.kyoto-u.ac.jp

Kentaro Inui
Graduate School of Information Science,
Nara Institute of Science and Technology
inui@is.naist.jp

Abstract

Towards deep analysis of compositional classes of paraphrases, we have examined a class-oriented framework for collecting paraphrase examples, in which sentential paraphrases are collected for each paraphrase class separately by means of automatic candidate generation and manual judgement. Our preliminary experiments on building a paraphrase corpus have so far been producing promising results, which we have evaluated according to cost-efficiency, exhaustiveness, and reliability.

1 Introduction

Paraphrases are alternative ways of conveying the same content. The technology for paraphrase generation and recognition has drawn the attention of an increasing number of researchers because of its potential contribution to a broad range of natural language applications.

Paraphrases can be viewed as monolingual translations. From this viewpoint, research on paraphrasing has adapted techniques fostered in the literature of machine translation (MT), such as transformation algorithms (Lavoie et al., 2000; Takahashi et al., 2001), corpus-based techniques for paraphrase pattern acquisition (Barzilay and McKeown, 2001; Shinyama and Sekine, 2003; Quirk et al., 2004), and fluency measurements (Lapata, 2001; Fujita et al., 2004).

One thing the paraphrasing community is still lacking is shared collections of paraphrase examples that could be used to analyze problems underlying the tasks and to evaluate the performance of systems under development. To our best knowledge, the paraphrase corpus developed by Dolan et al. (2004) is one of the very few collections available for free[1]. Development of paraphrase corpora raises several issues: what sorts of paraphrases should be collected, where paraphrase examples can be obtained from, how the coverage and quality of the corpus can be ensured, how manual annotation cost can be effectively reduced, and how collected examples should be organized and annotated.

Obviously these issues should be discussed with the purpose of each individual corpus taken into account. In this paper, we address the issues of building a gold-standard corpus that is to be used to evaluate paraphrase generation models and report on our preliminary experiences taking Japanese as a target language. Our approach is characterized by the following:

- We define a set of paraphrase classes based on the syntactic features of transformation patterns.
- We separately collect paraphrase examples for each paraphrase class that are considered to be linguistically explainable.
- We use a paraphrase generation system to exhaustively collect candidate paraphrases from a given text collection, which are then manually labeled.

[1] 5801 sentence pairs from their comparable corpus have been judged manually and available from http://research.microsoft.com/research/nlp/msr_paraphrase.htm

2 Goal

Paraphrases exhibit a wide variety of patterns ranging from lexical paraphrases to syntactic transformations and their combinations. Some of them are highly inferential or idiomatic and do not seem easy to generate only with syntactic and semantic knowledge. Such groups of paraphrases require us to pursue corpus-based acquisition methods such as those described in Section 3.

More importantly, however, we can also find quite a few patterns of paraphrases that exhibit a degree of regularity. Those groups of paraphrases have a potential to be compositionally explained by combining syntactic and semantic properties of their constituent words. For instance, the following paraphrases[2] in Japanese are considered to be of these groups.

(1) s. *eiga-ni* ***shigeki-o*** ***uke****-ta*.
film-DAT inspiration-ACC to receive-PAST
I received an inspiration from the film.
t. *eiga-ni* ***shigeki-s-are****-ta*.
film-DAT to inspire-PASS-PAST
I was inspired by the film.

(2) s. *sentakumono-****ga*** *soyokaze-****ni*** ***yureru***.
laundry-NOM breeze-DAT to sway-PRES
The laundry sways in the breeze.
t. *soyokaze-****ga*** *sentakumono-****o*** ***yurasu***.
breeze-NOM laundry-ACC to sway-PRES
The breeze makes the laundry sways.

(3) s. *glass-****ni*** *mizu-****o*** *mitashi-ta.*
glass-DAT water-ACC to fill-PAST
I filled water into the glass.
t. *glass-****o*** *mizu-****de*** *mitashi-ta*.
glass-ACC water-IMP to fill-PAST
I filled the glass with water.

(4) s. *kare-wa* ***kikai-sousa****-ga* ***jouzu****-da*.
he-TOP machine operation-NOM be good-PRES
He is good at machine operation.
t. *kare-wa* ***kikai****-o* ***jouzu****-ni* ***sousa-suru***.
he-TOP machine-ACC well-ADV to operate-PRES
He operates machines well.

(5) s. *heya-wa mou* ***atatamat-teiru***.
room-TOP already to be warmed-PERF
The room has already been warmed up.
t. *heya-wa mou* ***atatakai***.
room-TOP already be warm-PRES
The room is warm.

[2]For each example, "s" and "t" denote an original sentence and its paraphrase, respectively.

In example (1), a verb phrase, "*shigeki-o uke-ta* (to receive an inspiration)," is paraphrased into a verbalized form of the noun, "*shigeki-s-are-ta* (to be inspired)." We can find a number of paraphrases that exhibit a similar pattern of syntactic transformation in the same language and group such paraphrases into a single class, which is possibly labeled "*paraphrasing of light-verb construction*." Likewise, paraphrases exemplified by (2) constitute another class, so-called *transitivity alternation*. Example (3) is of the *locative alternation* class and example (4) the *compound noun decomposition* class. In example (5), a verb "*atatamaru* (to be warmed)" is paraphrased into its adjective form, "*atatakai* (be warm)." Paraphrases involving such a lexical derivation are also in our concern.

One can learn the existence of such groups of paraphrases and the regularity each group exhibits from the linguistic literature (Mel'čuk and Polguère, 1987; Jackendoff, 1990; Kageyama, 2001). According to Jackendoff and Kageyama, for instance, both transitivity alternation and locative alternation can be explained in terms of the syntactic and semantic properties of the verb involved, which are represented by what they call *Lexical Conceptual Structure*. The systematicity underlying such linguistic accounts is intriguing also from the engineering point of view as it could enable us to take a more theoretically motivated but still practical approach to paraphrase generation.

Aiming at this goal leads us to consider building a paraphrase corpus which enables us to evaluate paraphrase generation systems and conduct error analysis for each paraphrase class separately. Our paraphrase corpus should therefore be organized according to paraphrase classes. More specifically, we consider a paraphrase corpus such that:

- The corpus consists of a set of subcorpora.
- Each subcorpus is a collection of paraphrase sentence pairs of a paraphrase class.
- Paraphrases collected in a subcorpus sufficiently reflect the distribution of the occurrences in the real world.

Given a paraphrase class and a text collection, the goal of building a paraphrase corpus is to collect paraphrase examples belonging to the class

as exhaustively as possible from the text collection at a *minimal human labor cost*. The resultant corpus should also be *reliable*.

3 Related work

Previous work on building paraphrase corpus (collecting paraphrase examples) can be classified into two directions: manual production of paraphrases and automatic paraphrase acquisition.

3.1 Manual production of paraphrases

Manual production of paraphrase examples has been carried out in MT studies.

For example, Shirai et al. (2001) and Kinjo et al. (2003) use collections of Japanese-English translation sentence pairs. Given translation pairs, annotators are asked to produce new translations for each side of the languages. Sentences that have an identical translation are collected as equivalents, i.e., paraphrases. Shimohata (2004), on the other hand, takes a simpler approach in which he asks annotators to produce paraphrases of a given set of English sentences.

Obviously, if we simply asked human annotators to produce paraphrases of a given set of sentences, the labor cost would be expensive while the coverage not guaranteed. Previous work, however, has averted their eyes from evaluating the cost-efficiency of the method and the coverage of the collected paraphrases supposedly because their primary concern was to enhance MT systems.

3.2 Automatic paraphrase acquisition

Recently, paraphrase examples have been automatically collected as a source of acquiring paraphrase knowledge, such as pairs of synonymous phrases and syntactic transformation templates.

Some studies exploit topically related articles derived from multiple news sources (Barzilay and Lee, 2003; Shinyama and Sekine, 2003; Quirk et al., 2004; Dolan et al., 2004). Sentence pairs that are likely to be paraphrases are automatically collected from the parallel or comparable corpora, using such clues as overlaps of content words and named entities, syntactic similarity, and reference description, such as date of the article and positions of sentences in the articles.

Automatic acquisition from parallel or comparable corpora, possibly in combination with manual correction, could be more cost-efficient than manual production. However, it would not ensure coverage and quality, because sentence pairing algorithms virtually limit the range of obtainable paraphrases and products tend to be noisy.

Nevertheless, automatic methods are useful to discover a variety of paraphrases that need further exploration. We hope that our approach to corpus construction, which we present below, will work complementary to those directions of research.

4 Proposed method

Recall that we require a corpus that reflects the distribution of the occurrences of potential paraphrases of each class because we aim to use it for linguistic analysis and quantitative evaluation of paraphrase generation models.

Since the issues we address here are highly empirical, we need to empirically examine a range of possible methods to gain useful methodological insights. As an initial attempt, we have so far examined a simple method which falls in the middle of the aforementioned two approaches. The method makes use of an existing paraphrase generation system to reduce human labor cost as well as to ensure coverage and quality:

Step 1. For a given paraphrase class, develop a set of *morpho-syntactic paraphrasing patterns* and lexical resources.

Step 2. Apply the patterns to a given text collection using the paraphrasing system to generate a set of *candidate paraphrases*.

Step 3. Annotate each candidate paraphrase with information of the appropriateness according to a set of *judgement criteria*.

We use morpho-syntactic paraphrasing patterns derived from paraphrase samples in an analogous way to previous methods such as (Dras, 1999). For instance, from example (1), we derive a paraphrasing pattern for paraphrasing of light-verb constructions:

(6) s. $N\text{-}o_{(\Rightarrow V)}$ V
N-ACC V
t. $V(N)$
$V(N)$

where N is a variable which matches with a noun, V a verb, $V(N)$ denotes the verbalized form of

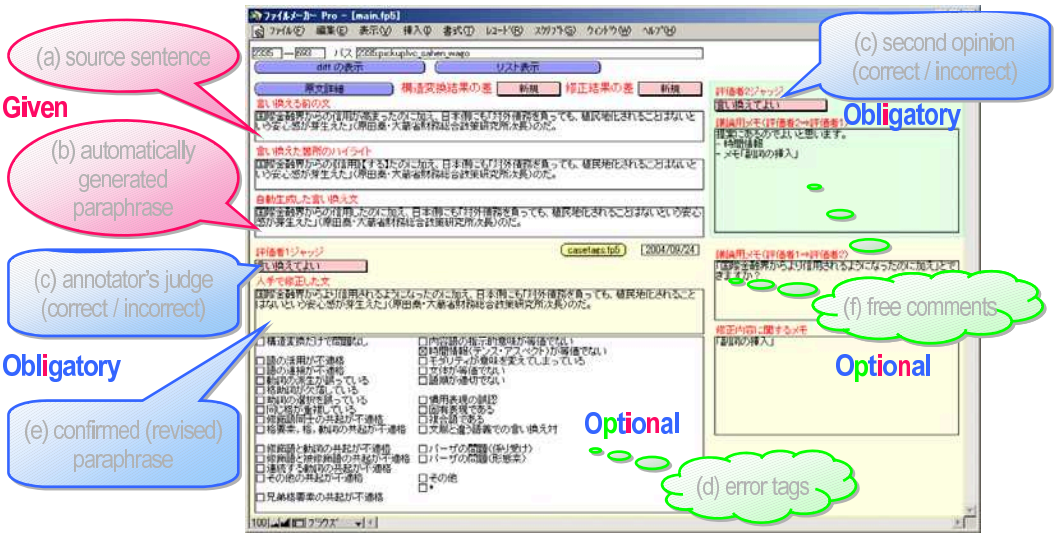

Figure 1: Annotation schema.

N, and the subscripted arrow in (6s) indicates that N-*o* depends on V.

To exhaustively collect paraphrase examples from a given text collection, we should not excessively constrain paraphrasing patterns. To avoid overly generating anomalies, on the other hand, we make use of several lexical resources. For instance, pairs of a deverbal noun and its transitive form are used to constrain N and $V(N)$ in pattern (6). This way, we combine syntactic transformation patterns with lexical constraints to specify a paraphrase class. This approach is practical given the recent advances of *shallow parsers*.

For the judgement on appropriateness in Step 3, we create a set of criteria separately for each paraphrase class. When the paraphrase class in focus is specified, the range of potential errors in candidate generation tends to be predictable. We therefore specify judgement criteria in terms of a typology of potential errors (Fujita and Inui, 2003); namely, we provide annotators with a set of conditions for ruling out inappropriate paraphrases.

Annotators judge each candidate paraphrase with a view of an RDB-based annotation tool (Figure 1). Given (a) a source sentence and (b) an automatically generated candidate paraphrase, human annotators are asked to (c) judge the appropriateness of it and, if it is inappropriate, they are also asked to (d) classify the underlying errors into a predefined taxonomy, and make (e) appropriate revisions (if possible) and (f) format-free comments.

5 Preliminary trials

To examine how the proposed method actually work regarding the issues, we conducted preliminary trials, taking two classes of Japanese paraphrases: *paraphrasing of light-verb constructions* and *transitivity alternation*. This section describes the settings for each paraphrase class.

We sampled a collection of source sentences from one year worth of newspaper articles: *Nihon Keizai Shinbun*[3], 2000, where the average sentence length was 25.3 words. The reason why we selected newspaper articles as a sample source was that most of the publicly available shallow parsers for Japanese were trained on a tree-bank sampled from newspaper articles, and a newspaper corpus was available in a considerably large scale. We used for candidate generation the morphological analyzer ChaSen[4], the dependency structure analyzer CaboCha[5], and the paraphrase generation system KURA[6].

Two native speakers of Japanese, adults graduated from university, were employed as annotators. The process of judging each candidate paraphrase is illustrated in Figure 2. The first annotator was asked to make judgements on each candidate paraphrase. The second annotator inspected all the candidates judged correct by the first an-

[3]http://sub.nikkeish.co.jp/gengo/zenbun.htm
[4]http://chasen.naist.jp/
[5]http://chasen.org/~taku/software/cabocha/
[6]http://cl.naist.jp/kura/doc/

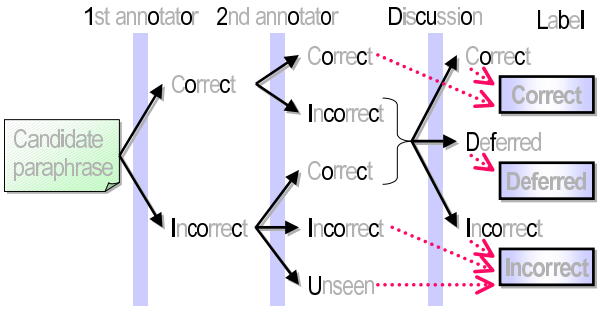

Figure 2: Judgement procedure.

notator. To reduce the labor cost, only a small subset of candidates that the first annotator judged incorrect were checked by the second annotator, leaving the rest labeled incorrect. Once in several days, the annotators discussed cases on which they disagreed, and if possible revised the annotation criteria. When the discussion did not reach a consensus, the judgement was deferred.

5.1 Paraphrasing of light-verb constructions (LVC)

An example of this class is given in (1). A light-verb construction consists of a deverbal noun ("*shigeki* (inspiration)" in example (1)) governed by a light-verb ("*ukeru* (to receive)"). A paraphrase of this class is a pair of a light-verb construction and its unmarked form, which consists of the verbalized form of the deverbal noun where the light-verb is removed.

Let N, V be a deverbal noun and a verb, and $V(N)$ be the verbalized form of N. Paraphrases of this class can be represented by the following paraphrasing pattern:

(7) s. N-{*ga, o, ni*}$_{(\Rightarrow V)}$ V
N-{NOM, ACC, DAT} V
t. $V(N)$
$V(N)$

In the experiment, we used three more patterns to gain the coverage.

We then extracted 20,155 pairs of deverbal noun and its verbalized form (e.g. "*shigeki* (inspiration)" and "*shigeki-suru* (to inspire)") from the Japanese word dictionary, IPADIC (version 2.6.3)[3]. This set was used as a restriction on nouns that can match with N in a paraphrasing pattern. On the other hand, we made no restriction on V, because we had no exhaustive list of light-verbs. The patterns were automatically compiled into pairs of dependency trees with uninstantiated components, and were applied to source sentences with the paraphrase generation system, which carried out dependency structure-based pattern matching. 2,566 candidate paraphrases were generated from 10,000 source sentences.

In the judgement phase, the annotators were also asked to revise erroneous candidates if possible. The following revision operations were allowed for LVC:

- Change of conjugations
- Change of case markers
- Insert adverbs
- Append verbal suffixes, such as voice, aspect, or mood devices

When pattern (7) is applied to sentence (1s), for instance, we need to add a voice device, "*are* (passive)," to correctly produce (1t). In example (8), on the other hand, an aspectual device, "*dasu* (inchoative)," is appended, and a case marker, "*no* (GEN)," is replaced with "*o* (ACC)."

(8) s. *concert-no ticket-**no** **hanbai-o** **hajime**-ta.*
concert-GEN ticket-GEN sale-ACC to start-PAST
We started to sale tickets for concerts.
t. *concert-no ticket-**o** **hanbai-shi-dashi**-ta.*
concert-GEN ticket-ACC to sell-INCHOATIVE-PAST
We started selling tickets for concerts.

So far, 1,114 candidates have been judged[7] with agreements on 1,067 candidates, and 591 paraphrase examples have been collected.

5.2 Transitivity alternation (TransAlt)

This class of paraphrases requires a collection of pairs of intransitive and transitive verbs, such as "*yureru* (to sway)" and "*yurasu* (to sway)" in example (2). Since there was no available resource of such knowledge, we newly created a minimal set of intransitive-transitive pairs that were required to cover all the verbs appearing in the source sentence set (25,000 sentences). We first retrieved all the verbs from the source sentences using a set of extraction patterns implemented in the same manner as paraphrasing patterns. Example (9) is one of the patterns used, where N_x matches with a noun, and V a verb.

[7] 983 candidates for the first 4,500 sentences were fully judged, and 131 candidates were randomly sampled from the remaining portion.

(9) s. N_1-*ga*$_{(\Rightarrow V)}$ N_2-*ni*$_{(\Rightarrow V)}$ V
N_1-NOM N_2-DAT V
t. no change.

We then manually examined the transitivity of each of 800 verbs that matched with V, and collected 212 pairs of intransitive verb v_i and its transitive form v_t. Using them as constraints, we implemented eight paraphrasing patterns as in (10).

(10) s. N_1-*ga*$_{(\Rightarrow V_i)}$ N_2-*ni*$_{(\Rightarrow V_i)}$ V_i
N_1-NOM N_2-DAT V_i
t. N_2-*ga*$_{(\Rightarrow V_t(V_i))}$ N_1-*o*$_{(\Rightarrow V_t(V_i))}$ $V_t(V_i)$
N_2-NOM N_1-ACC $V_t(V_i)$

where V_i and $V_t(V_i)$ are variables that match with v_i and v_t, respectively. By applying the patterns to the same set of source sentences, we obtained 985 candidate paraphrases.

We created a set of criteria for judging appropriateness (an example will be given in Section 6.4) and revision examples for the following operations allowed for this trial:

- Change of conjugations
- Change of case markers
- Change of voices

964 candidates have gained an agreement, and 484 paraphrase examples have been collected.

6 Results and discussion

Table 1 gives some statistics of the resultant paraphrase corpora. Figures 3 and 4 show the number of candidate paraphrases, where the horizontal axes denote the total working hours of two annotators, and the vertical axes the number of candidate paraphrases. The numbers of judged, correct, incorrect, and deferred candidates are shown.

6.1 Efficiency

2,031 candidate paraphrases have so far been judged in total and 1,075 paraphrase examples have been collected in 287.5 hours. The judgement was performed at a constant pace: 7.1 candidates (3.7 examples) in one hour. It is hard to compare these results with other work because no previous study quantitatively evaluate the efficiency in terms of manual annotation cost. However, we feel that the results have so far been satisfiable.

For each candidate paraphrase judged incorrect, the annotators were asked to classify the underlying errors into the fixed error types ((d) in Figure 1). This error classification consumed extra time because it required linguistic expertise which the annotators were not familiar with.

Table 1: Statistics of the resultant corpora.

Paraphrase class	LVC	TransAlt
# of source sentences	10,000	25,000
# of patterns	4	8
Type of lexical resources	$\langle n, v_n \rangle$	$\langle v_i, v_t \rangle$
Size of lexical resource	20,155	212
# of candidates	2,566	985
# of judged candidates	1,067	964
# of incorrect candidates	520	503
# of correct candidates	547	461
# of paraphrase examples	591	484
Working hours	118	169.5

TransAlt was 1.75 times more time-consuming than LVC because the definition of TransAlt involved several delicate issues, which made the judgement process complicated. We return to this issue in Section 6.4.

6.2 Exhaustiveness

To estimate how exhaustively the proposed method collected paraphrase examples, we randomly sampled 750 sentences from the 4,500 sentences that were used in the trial for LVC, and manually checked whether the LVC paraphrasing could apply to each of them. As a result, 206 examples were obtained, 158 of which were those already collected by the proposed method. Thus, the estimated exhaustiveness was 77% (158 / 206). Our manual investigation into the missed examples has revealed that 47 misses could have been automatically generated by enhancing paraphrasing patterns and dictionaries, while only one example was missed due to an error in shallow parsing. 34 cases of the 48 misses could have been collected by adding a couple of paraphrasing patterns. For example, pattern (11) verbalizes a noun followed by a nominalizing suffix, "*ka* (-ize)," as in (12).

(11) s. N-*ka*-{*ga, o, ni*}$_{(\Rightarrow V)}$ V
N-ize-{NOM, ACC, DAT} V
t. $V(N$-*ka*$)$
$V(N$-ize$)$

(12) s. *kore-wa* *kin'yu-shijo-**no*** ***kassei-ka-ni***
this-TOP financial market-GEN activation-DAT
muke-ta *kisei-kanwa-saku-da*.
to address-PAST deregulation plan-COP
This is a deregulation plan aiming at the activation of financial market.

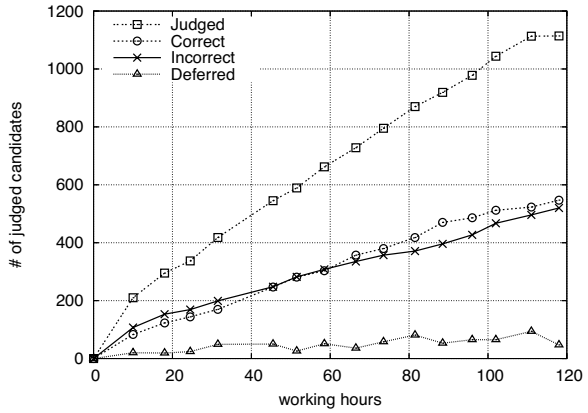

Figure 3: # of judged candidates (LVC).

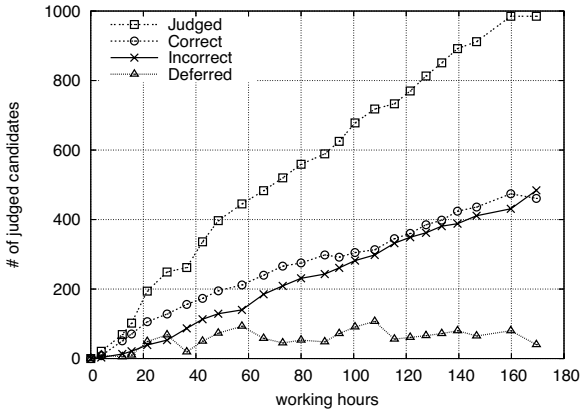

Figure 4: # of judged candidates (TransAlt).

t. *kore-wa kin'yu-shijo-**o***
this-TOP financial market-ACC
***kassei-ka-suru** kisei-kanwa-saku-da.*
to activate-PRES deregulation plan-COP
This is a deregulation plan which activates financial market.

We cannot know if we have adequate paraphrasing patterns and resources before trials. Therefore, manual examination is necessary to refine them to bridge gap between the range of paraphrases that can be automatically generated and those of the specific class we consider.

6.3 Reliability

Ideally, more annotators should be employed to ensure the reliability of the products, which, however, leads to a matter of balancing the trade-off. Instead, we specified the detailed judgement criteria for each paraphrase class, and asked the annotators to reconsider marginal cases several days later and to make a discussion when judgements disagreed. The agreement ratio for correct candidates between two annotators increased as they became used to the task. In the trial for LVC, for example, the agreement ratio for each day changed from 74% (day 3) to 77% (day 6), 88% (day 9), and 93% (day 11). This indicates that the judgement criteria were effectively refined based on the feedback from inter-annotator discussions on marginal and disagreed cases. To evaluate the reliability of our judgement procedure more precisely, we are planing to employ the third annotator who will be asked to judge all the cases independently of the others.

6.4 How we define paraphrase classes

One of the motivations behind our class-based approach is an expectation that specifying the target classes of paraphrases would simplify the awkward problem of defining the boundary between paraphrases an non-paraphrases. Our trials for the two paraphrase classes, however, have revealed that it can still be difficult to create a clear criterion for judgement even when the paraphrase class in focus is specified.

As one of the criteria for TransAlt, we tested the agentivity of the nominative case of intransitive verbs. The test used an adverb, "*muzukara* (by itself)," and classified a candidate paraphrase as incorrect if the adverb could be inserted immediately before the intransitive verb. For example, we considered example (13) as a correct paraphrase of the TransAlt class whereas (14) incorrect because the agentivity exhibited by (14s) did not remain in (14t).

(13) s. *kare-ga soup-**o** **atatame**-ta.*
he-NOM soup-ACC to warm up-PAST
He warmed the soup up.
t. *soup-**ga** **atatamat**-ta.* **(correct)**
soup-NOM to be warmed up-PAST
The soup was warmed up (by somebody).

(14) s. *kare-ga koori-**o** **tokashi**-ta.*
he-NOM ice-ACC to melt (vt)-PAST
He melted the ice.
t. *koori-**ga** **toke**-ta.* **(incorrect)**
ice-NOM to melt (vi)-PAST
The ice melted (by itself).

However, one might regard both paraphrases incorrect because the information given by the nominative argument of the source sentence is

dropped in the target in both cases. Thus, the problem still remains. Nevertheless, our approach will provide us with a considerable amounts of concrete data, which we hope will lead us to better understanding of the issue.

7 Conclusion

Towards deep analysis of compositional classes of paraphrases, we have examined a class-oriented framework for collecting paraphrase examples, in which sentential paraphrases are collected for each paraphrase class separately by means of automatic candidate generation and manual judgement. Our preliminary experiments on building a paraphrase corpus have so far been producing promising results, which we have evaluated according to cost-efficiency, exhaustiveness, and reliability. The resultant corpus and resources will be available for free shortly. Our next step is directed to targeting a wider range of paraphrase classes.

References

Regina Barzilay and Kathleen R. McKeown. 2001. Extracting paraphrases from a parallel corpus. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics* (*ACL*), pages 50–57.

Regina Barzilay and Lillian Lee. 2003. Learning to paraphrase: an unsupervised approach using multiple-sequence alignment. In *Proceedings of the 2003 Human Language Technology Conference and the North American Chapter of the Association for Computational Linguistics* (*HLT-NAACL*), pages 16–23.

Bill Dolan, Chris Quirk, and Chris Brockett. 2004. Unsupervised construction of large paraphrase corpora: exploiting massively parallel news sources. In *Proceedings of the 20th International Conference on Computational Linguistics* (*COLING*), pages 350–356.

Mark Dras. 1999. *Tree adjoining grammar and the reluctant paraphrasing of text*. Ph.D. thesis, Division of Information and Communication Science, Macquarie University.

Atsushi Fujita and Kentaro Inui. 2003. Exploring transfer errors in lexical and structural paraphrasing. *IPSJ Journal*, 44(11):2826–2838. (in Japanese).

Atsushi Fujita, Kentaro Inui, and Yuji Matsumoto. 2004. Detection of incorrect case assignments in automatically generated paraphrases of Japanese sentences. In *Proceedings of the 1st International Joint Conference on Natural Language Processing* (*IJCNLP*), pages 14–21.

Ray Jackendoff. 1990. *Semantic structures*. The MIT Press.

Taro Kageyama, editor. 2001. *Semantics and syntax of verb: comparable study between Japanese and English*. Taishukan Shoten. (in Japanese).

Yumiko Kinjo, Kunio Aono, Keishi Yasuda, Toshiyuki Takezawa, and Genichiro Kikui. 2003. Collection of Japanese paraphrases of basic expressions on travel conversation. In *Proceedings of the 9th Annual Meeting of the Association for Natural Language Processing*, pages 101–104. (in Japanese).

Maria Lapata. 2001. A corpus-based account of regular polysemy: the case of context-sensitive adjectives. In *Proceedings of the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics* (*NAACL*), pages 63–70.

Benoit Lavoie, Richard Kittredge, Tanya Korelsky, and Owen Rambow. 2000. A framework for MT and multilingual NLG systems based on uniform lexico-structural processing. In *Proceedings of the 6th Applied Natural Language Processing Conference and the 1st Meeting of the North American Chapter of the Association for Computational Linguistics* (*ANLP-NAACL*), pages 60–67.

Igor Mel'čuk and Alain Polguère. 1987. A formal lexicon in meaning-text theory (or how to do lexica with words). *Computational Linguistics*, 13(3-4):261–275.

Chris Quirk, Chris Brockett, and William Dolan. 2004. Monolingual machine translation for paraphrase generation. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing* (*EMNLP*), pages 142–149.

Mitsuo Shimohata. 2004. *Acquiring paraphrases from corpora and its application to machine translation*. Ph.D. thesis, Graduate School of Information Science, Nara Institute of Science and Technology.

Yusuke Shinyama and Satoshi Sekine. 2003. Paraphrase acquisition for information extraction. In *Proceedings of the 2nd International Workshop on Paraphrasing: Paraphrase Acquisition and Applications* (*IWP*), pages 65–71.

Satoshi Shirai, Kazuhide Yamamoto, and Francis Bond. 2001. Japanese-English paraphrase corpus. In *Proceedings of the 6th Natural Language Processing Pacific Rim Symposium* (*NLPRS*) *Workshop on Language Resources in Asia*, pages 23–30.

Tetsuro Takahashi, Tomoya Iwakura, Ryu Iida, Atsushi Fujita, and Kentaro Inui. 2001. KURA: a transfer-based lexico-structural paraphrasing engine. In *Proceedings of the 6th Natural Language Processing Pacific Rim Symposium* (*NLPRS*) *Workshop on Automatic Paraphrasing: Theories and Applications*, pages 37–46.