

# SMT reranked NMT

Terumasa EHARA

Ehara NLP Research Laboratory  
Seijo, Setagaya, Tokyo, JAPAN

eharate @ gmail . com

## Abstract

System architecture, experimental settings and experimental results of the EHR team for the WAT2017 tasks are described. We participate in three tasks: JPCen-ja, JPCzh-ja and JPCko-ja. Although the basic architecture of our system is NMT, reranking technique is conducted using SMT results. One of the major drawback of NMT is under-translation and over-translation. On the other hand, SMT infrequently makes such translations. So, using reranking of n-best NMT outputs by the SMT output, discarding such translations can be expected. We can improve BLEU score from 46.03 to 47.08 by this technique in JPCzh-ja task.

## 1 Introduction

Rapidly progressing of NMT techniques make paradigm change in machine translation not only for the research purpose but for the practical field. Although the NMT provides high quality and fluent translations, it has several drawbacks. One of them is under- and over-translation which is infrequent in a SMT output.

We propose a reranking method for n-best NMT outputs using a SMT output. We compare n-best NMT outputs with a SMT output by the measure of IMPACT (Echizen-ya and Araki, 2007) which is one of the automatic evaluation measure of machine translation results. The NMT output which has the highest IMPACT score referring to SMT output is selected as the system output.

In the following sections, we describe system architecture and experimental settings in section 2, experimental results and discussions in section 3 and conclusion in section 4.

## 2 System architecture and experimental settings

### 2.1 Overall system architecture

Our system architecture is shown in Figure 1. An input source sentence is fed to the NMT part and also to the SMT part. NMT part outputs n-best translations (“NMT translation 1” to “NMT translation n”) and SMT part outputs another translation (“SMT translation”). Reranking part compares NMT translations with SMT translation and reranks them. The best reranked “NMT translation i” is outputted.

### 2.2 NMT part

We use OpenNMT (Minh-Thang Luong et al., 2015) in NMT part.

Segmentation of English sentences is sub word based. The English segmenter segments each non-alphabetical characters (characters except for A to Z and a to z) as separate words. Segmentation of Chinese sentences and Korean sentences are both word based and character based. Word segmentation policy for these languages are described in the previous paper (Ehara, 2016). Japanese segmentation is word based, sub word based and character based. For JPCzh-ja task and JPCko-ja task, word based and character based Japanese segmenters are used. The word based Japanese segmenters are described in the previous paper (Ehara, 2016). For JPCen-ja task, we use sub word based Japanese segmenter which segments each special characters (characters except for Hiragana, Kanji, Katakana and Roman characters) as separate words, in addition to Juman’s word segmentation (Kurohashi et al., 1994).

Option settings for OpenNMT are as follows:  
Source sequence length (-src\_seq\_len): 100 (word based), 120 (sub word based), 250 (character based); Target sequence length (-tgt\_seq\_len):

100 (word based), 120 (sub word based), 250 (character based); Encoder type (-encoder\_type): brnn (bidirectional recurrent NN); Replace unknown word (-replace\_unk): yes; Unknown word dictionary (-phrase\_table): yes (see 2.3); Beam size (-beam\_size): 50; N-best size (-n\_best): 50.

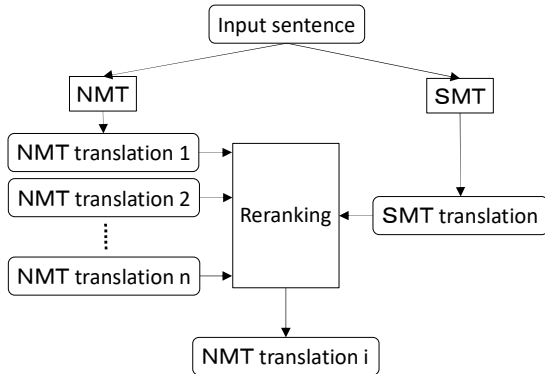


Figure 1: System architecture

### 2.3 SMT part

Our SMT system is phrase-based SMT by Moses v.3 (Koehn et al., 2003) with default option settings. For JPCen-ja task and JPCzh-ja task, pre-ordering is applied. The preordering system is same as described in the previous papers (Ehara, 2015; Ehara, 2016).

We use unknown dictionary for NMT part. It is made from the phrase-table of Moses. For every source word, we select the target phrase which has the highest translation probability for the source word. And the unknown word dictionary is constructed as the source word and target phrase pairs.

### 2.4 Reranking part

For reranking of n-best outputs of NMT part, we use automatic evaluation measure IMPACT

(Echizen-ya and Araki, 2007). For the preliminary study, we compared BLEU, RIBES and IMPACT with human evaluation score JPO adequacy by the WAT2016’s evaluation results (Nakazawa et al., 2016). As the results, we found IMPACT was the best correlated score with JPO adequacy. Then we use IMPACT as the reranking measure. Reranking part calculates IMPACT score for NMT’s n-best translations with SMT translation as the reference. And the best translation which has the highest IMPACT score is outputted as the system output.

## 3 Experimental results and discussions

The official evaluation results of our submissions are shown in Table 1 (Nakazawa et al., 2017). In the Table 1, “Original system” means the NMT without reranking and “SMT” means SMT part of our system.

For JPCen-ja task, reranking decreases BLEU, RIBES and AMFM scores and also HUMAN score. Although the overall evaluation result doesn’t show the effectiveness of the reranking, several improvements are observed. Examples are listed in Table 2. Original translation of the example 1 has under-translation. Only the first two words (The oldest) and the punctuation mark (.) are translated in the original translation. Original translations of example 2 has also under-translation. None of words “( ACT , READ , PRE ) , GBSTB , GBSTT , FXb 2 , PUMP , FXB , FXT , SWL , and RFX” is translated. On the other hand, reranking system does not make such under-translations. Original translation of example 3 has over-translation. “異なる (differ)” occurs two times. But the reranked translation has no over-translation.

Task	Data ID	System	Segment.	BLEU	RIBES	AMFM	HUMAN	JPO adeq.
JPCen-ja	1406	Reranking	Subword	44.44	0.8610	0.7471	58.250	----
	1407	Original	Subword	44.63	0.8667	0.7478	60.000	4.63
	----	SMT	Word	36.20	0.8128	0.7237	----	----
JPCzh-ja	1408	Reranking	Word	47.08	0.8591	0.7564	68.250	----
	1415	Original	Word	46.03	0.8586	0.7559	----	----
	1414	Reranking	Character	46.52	0.8596	0.7614	69.750	4.31
	1409	Original	Character	45.27	0.8544	0.7571	----	----
	----	SMT	Word	40.79	0.8270	0.7384	----	----
JPCko-ja	1416	Reranking	Word	71.52	0.9445	0.8661	6.250	----
	1418	Original	Word	70.23	0.9432	0.8623	----	----
	1417	Reranking	Character	71.36	0.9461	0.8711	11.250	4.81
	1419	Original	Character	69.42	0.9364	0.8605	----	----
	----	SMT	Word	71.08	0.9440	0.8645	----	----

Table 1: Official evaluation results (Japanese segmenter is Juman)

1	source	The oldest is the capacitive divider bridge , one arm of which consists of a reference capacitance and the other of the capacitance to be measured .
	reference	最も古いものは容量分圧器ブリッジであり、その1つのアームは基準静電容量から成り、その他の静電容量が測定される。
	reranking	最も古いものは容量分圧ブリッジであり、その1つのアームは、測定されるべき基準キャパシタンスと他のキャパシタンスとから成る。
	original	最も古い。
	SMT	最も古いものは、容量性ディバイダブリッジ、基準キャパシタンスと測定されるべきキャパシタンスの他の1つを構成するアームである。
2	source	FIG . 5 shows a portion of waveforms of a simulation result of the circuit in FIG . 4 , and commands ( ACT , READ , PRE ) , GBSTB , GBSTT , FXb 2 , PUMP , FXB , FXT , SWL , and RFX are shown as signals .
	reference	図5には、図4の回路のシミュレーション結果の波形が一部示されており、信号として、コマンド ( ACT 、 READ 、 PRE ) 、 GBSTB 、 GBSTT 、 FXb 2 、 PUMP 、 FXB 、 FXT 、 SWL 、 RFX が示されている。
	reranking	図5は、図4の回路のシミュレーション結果の一部を示すものであり、コマンド ( ACT 、 READ 、 PRE ) 、 GBSTB 、 GBSTT 、 FXb 2 、 PUMP 、 FXB 、 SWL 、 RFX を信号として示す。
	original	図4の回路のシミュレーション結果の一部と、図4の回路のシミュレーション結果の一部を示す図である。
	SMT	図5は、図4の回路のシミュレーション結果の波形の一部を示す、及びコマンド ( ACT 、 READ 、 PRE ) 、 GBSTB 、 GBSTT 、 FXb 2 、 PUMP 、 FXB 、 FXT RFX SWLとは、信号として示されている。
3	source	However, specifications sometimes differ from one image forming apparatus to another.
	reference	しかし、画像形成装置ごとにスペックが異なる場合がある。
	reranking	しかし、ある仕様は、時々、1つの画像形成装置とは異なる。
	original	しかし、時々、1つの画像形成装置とは異なる仕様が異なる。
	SMT	しかし、仕様は、装置を形成する1つの画像から別の画像へ時には異なる。

Table 2: JPCen-ja task’s examples having effectiveness of reranking (recover of under-translation: example 1 and 2; recover of over-translation: example 3)

For JPCen-ja task, comparing our submission of data ID 1407 (EHR) and another submission (OTHER), BLEU score of the EHR is 44.63 and it is less than the OTHER’s score (50.27). On the other hand, HUMAN score of the EHR is 60.00 and it is greater than the OTHER’s score (56.25). There are 20 data that the BLEU<sup>1</sup> score of EHR is less than the OTHER’s score but the HUMAN score of EHR is greater than the OTHER’s score<sup>2</sup>.

We examine these data and find that several data have the differences between the source expression and the reference expression. Table 3

shows examples of source, reference, EHR output and OTHER output. Example 1 has the voice change (the source is passive and the reference is active). Example 2 has the topic change (the topic of the source is “valve” and the topic of the reference is “経路 (passage)”). Example 3 has the additive translation<sup>3</sup> (the source “GELD” corresponds the reference “有機ELディスプレイ (OELD, organic electroluminescent display)”). Example 4 has the subtractive translation<sup>4</sup> (the source has

<sup>1</sup> Sentence level BLEU is calculated by mteval-v13a.pl in the Moses package.

<sup>2</sup> For the BLEU score, “less” means “less or equal -10” and “greater” means “greater or equal 10”. If the difference of BLEU is between -10 to 10, it is considered “tie”.

<sup>3</sup> We distinguish between “additive translation” and “over-translation”. The former means the translation including

complemental information and the latter means the translation including needless information.

<sup>4</sup> We distinguish between “subtractive translation” and “under-translation”. The former means the translation omitting complemental information and the latter means the translation omitting needful information.

“given within this range” and the reference has no such expression).

Comparing JPO adequacy score of EHR and OTHER, the EHR’s score 4.63 is less than the OTHER’s score 4.75. Although HUMAN and JPO adequacy are both human evaluation, they

have a contradiction. One possible reason is that the evaluators of HUMAN score do not look at the reference translations but the evaluators of JPO adequacy score can look at the reference translations.

1	source	In FIG. 6 , the same symbols are used to the same elements as the elements shown in FIGS. 1-5
	reference	図 6 において、図 1 ~ 図 5 に示す要素と同じ要素には同じ符号を付してある。
	EHR	図 6 では、同じ記号が、図 1 ~ 5 に示される要素と同じ要素に使用される。
	OTHER	図 6 において、図 1 ~ 図 5 と同じ要素には同じ符号を付している。
2	source	A sixth on-off valve 202 is disposed on the fourth bypass passage 204 .
	reference	第 4 バイパス 経路 204 には 第 6 開閉 弁 202 が 設けられて いる 。
	EHR	第 6 の オン オフ 弁 202 は、第 4 バイパス 通路 204 上 に 配置 される 。
	OTHER	第 4 の バイパス 通路 204 には、第 6 の 開閉 弁 202 が 配 置 されて いる 。
3	source	The display device 4 is, for example, an LCD and an GELD.
	reference	表示 装置 4 は、例えば、液晶 ディスプレイ ( LCD , liquidcry staldisplay ) や 有機 EL ディスプレイ ( OELD , organicelectro - luminescentdisplay ) 等 である 。
	EHR	表示 装置 4 は、例えば LCD および GELD である 。
	OTHER	表示 装置 4 は、例えば LCD や OELD ( OELD , organicelectro - luminescentdisplay ) 等 である 。
4	source	In this case, the proportion of the additive given within this range corresponds to 3% or lower.
	reference	この 場合 の 添加 剤 の 添加 量 は 3 % 以下 である 。
	EHR	この 場合、この 範囲 内 に 与えらるる 添加 剤 の 割合 は、3 % 以下 に 相当 する 。
	OTHER	この 場合、添加 剤 の 割合 は、3 % 以下 である 。

Table 3: Different expressions between sources and references

(example 1: voice change; example 2: topic change; example 3: additive translation; example 4: subtractive translation)

For JPCzh-ja and JPCko-ja tasks, reranking increases BLEU, RIBES and AMFM scores. However, we don’t have a HUMAN scores comparing the reranking and the original for these tasks. Examples having the effectiveness of the reranking for these tasks are shown in Table 4 and Table 5.

Example 1 and 2 of Table 4 have under-translation in original translation. Example 3 of Table 4 has over-translation in original translation. Example 1 and 2 of Table 5 have under-translation in original translation.

1	source	图 3 ( A ) 是对 进行 2 次 通过 间隙 的 处理 的 高 分子 组 成 物 的 制 造 装 置 从 上 面 透 视 装 置 内 部 时 的 概 略 透 视 图 ， 图 3 ( B ) 是 图 3 ( A ) 的 装 置 的 P - Q 截 面 上 的 概 略 截 面 图 。
	reference	图 3 ( A ) は 、 間 隙 通 過 处 理 を 2 回 行 う 高 分 子 組 成 物 の 製 造 装 置 に つ い て 上 面 か ら 装 置 内 部 を 透 視 し た と き の 概 略 透 視 図 で あり 、 图 3 ( B ) は 、 图 3 ( A ) の 装 置 の P - Q 断 面 に お け る 概 略 断 面 図 で あり 。
	reranking	图 3 ( A ) は 、 ギ ャ ッ プ を 2 回 行 う 处 理 を 行 う 高 分 子 組 成 物 の 製 造 装 置 を 上 面 透 視 装 置 内 部 か ら 見 た 場 合 の 概 略 透 視 図 で あり 、 图 3 ( B ) は 、 图 3 ( A ) の 装 置 の P - Q 断 面 に お け る 概 略 断 面 図 で あり 。
	original	图 3 ( A ) は 、 图 3 ( A ) の 装 置 の P - Q 断 面 上 の 概 略 的 な 断 面 図 で あり 。
	SMT	图 3 ( A ) に つ い て は 、 2 次 ギ ャ ッ プ の 处 理 に よ っ て 高 分 子 組 成 物 の 製 造 装 置 を 上 面 か ら 装 置 内 部 透 視 時 の 概 略 斜 視 図 で あり 、 图 3 ( B ) は 、 图 3 ( A ) の 装 置 の P - Q 断 面 に お け る 概 略 断 面 図 で あり 。
2	source	此外 ， 根 据 元 素 的 话 ， 一 个 元 素 有 可 能 形 成 不 同 化 合 价 的 氧 化 物 。
	reference	また 、 元 素 に よ っ て は 、 1 つ の 元 素 が 異 な る 価 数 の 酸 化 物 を 形 成 す る こ と が 可 能 で あり 。
	reranking	また 、 元 素 に よ っ て は 、 1 つ の 元 素 が 異 な る 種 類 の 酸 化 物 を 形 成 す る こ と が 可 能 で あり 。
	original	また 、 元 素 に よ っ て は 、 異 な る 元 素 を 形 成 す る こ と が 可 能 で あり 。
	SMT	また 、 元 素 と に よ っ て は 、 異 な る 原 子 価 元 素 の 酸 化 物 を 形 成 す る こ と が 可 能 で あり 。
3	source	实 施 例 14
	reference	【 实 施 例 14 】
	reranking	实 施 例 14 :
	original	( 实 施 例 14 ) 实 施 例 14 に つ い て 説 明 す る 。
	SMT	实 施 例 14

Table 4: JPCzh-ja task's examples having effectiveness of reranking (recover of under-translation: example 1 and 2; recover of over-translation: example 3)

1	source	이어서, 용매로서 DINP ( 디 이소노닐 프탈 레이트 ) 183 질 량부 를 첨가 하였다 .
	reference	次いで、溶媒としてDINP ( ジイソニルフタレート ) 183 質量部 を添加した。
	reranking	次いで、溶媒としてDINP ( ジメチルホルムアミド ) 183 質量部 を添加した。
	original	次に、溶媒としてDINP (、) 183 質量部 を添加した。
	SMT	次いで、溶媒としてDINP (記述イソノアニールフタレート) 183質量部を添加した。
2	source	저장 장치 ( 70 ) 는 다음 을 포함 할 수 있 다 ( 도 15 및 도 16 ) : - 하나 혹은 그 이상 의 전기 배터리 ( 81 ) ; 또는 - 하이브리드 배터리 ( 82 ) 및 , 상 기 하이브리드 배터리 ( 82 ) 와 유효 하 게 연결 된 내연 기관 ( 83 ) .
	reference	貯蔵装置70は下記を備えることができる(図15と16): - 1台または2台以上の電気バッテリー81;または、- ハイブリッドバッテリー82と、前記ハイブリッドバッテリー82に接続されている内燃機関83。
	reranking	記憶装置70は、以下を含むことができる(図15及び図16): 1つまたはそれ以上の電気バッテリー81、又はハイブリッドバッテリー82、及び、ハイブリッドバッテリー82と有効に連結された内燃機関83。
	original	記憶装置70は、以下を含むことができる(図15および図16)。
	SMT	記憶装置70は以下を含むことができる(図15及び図16:-1つあるいはそれ以上の電気バッテリー81;又は-ハイブリッドバッテリー82及び、上記ハイブリッドバッテリー82と有効に連結された内燃機関(83)。

Table 5: JPCko-ja task’s examples having effectiveness of reranking (recover of under-translation: example 1 and 2)

For JPCzh-ja and JPCko-ja tasks, the word based translations have higher BLEU, RIBES and AMFM compared with the character based translations. However, HUMAN score of the word based translations are lower than the score of the character based translations.

For JPCzh-ja task, there are 7 data that the BLEU score of the word based translation is greater than the character based translation’s score but the HUMAN score of the word based translation is less than the character based translation’s score. Examples of such translations are listed in Table 6. Example 1 has an under-translation in the word based translation (“滴度”). Example 2 also has an under-translation in the word based translation (“実行される(进行的)”). Example 3 has miss translations both in the character based translation and the word based translation. In the character based translation, “取付面図” is used instead of “実装面図(安装面图)”. And in the word based translation, “分波器モジュール” is used instead of “デュプレクサモジュール(双工器模块)”. However, the latter miss translation is more significant than the former. Example 4 has another different translations. Character based

translation uses “が良い(好)”, and word based translation uses “に優れる”.

For JPCko-ja task, there are 3 data that the BLEU score of the word based translation is greater than the character based translation’s score and the HUMAN score of the word based translation is less than the character based translation’s score. Examples of such translations are listed in Table 7. Example 1 has a different translation. Literal translation of “연결” is “連結” and non-literal translation is “接続”. Example 2 shows the effectiveness of the unknown word translation in the character based translation. The expression

“디펜타에리트릴펜타아크릴레이트와 디펜타에리트릴” does not be translated in the word based translation. Example 3 has different translations. Character based translation uses “ブレーキ(브레이크)”, and word based translation uses “ブレーク”.

1	source	这进一步提示高估了CAZ028单价散装液 <b>滴度</b> 。
	reference	このことは、CAZ028一価バルクの力価が高く見積もられていることをさらに示唆する。
	char. based	これはさらに、CAZ028単価の分散液の <b>滴度</b> を高めることを示唆する。
	word based	これにより、CAZ028の一価ばら積みをさらに示唆する。
2	source	图9示出了电台110为释放额外的无线资源所 <b>进行的</b> 处理过程900的设计。
	reference	図9は、余分な無線リソースを放棄するために <b>局110</b> によって <b>実行される</b> 方法900の設計を示す。
	char. based	図9は、 <b>局110</b> が、追加の無線リソースを解放するために <b>行われる</b> 処理プロセス900の設計を示す。
	word based	図9は、追加の無線リソースを解放するために <b>局110</b> が処理プロセス900の設計を示す。
3	source	图4(A)是双工器模块的简要等效电路图, 图4(B)是 <b>双工器模块的安装面图</b> 。
	reference	図4(A)はデュプレクサモジュールの概略の等価回路図であり、図4(B)は <b>デュプレクサモジュールの実装面図</b> である。
	char. based	図4(A)は、デュプレクサモジュールの概略等価回路図であり、図4(B)は、 <b>デュプレクサモジュールの取付面図</b> である。
	word based	図4(A)はデュプレクサモジュールの概略的な等価回路図であり、図4(B)は <b>分波器モジュールの実装面図</b> である。
4	source	另一方面, 如果大于4.2倍, 则虽然耐水解性好, 但基材层(B)的凝聚强度降低, 因此不理想。
	reference	他方、4.2倍を超えると耐加水分解性は <b>良い</b> が基材層(B)の凝集強度が低くなり好ましくない。
	char. based	一方、4.2倍を超えると、耐加水分解性が <b>良い</b> が、基材層(B)の凝集強度が低下するため好ましくない。
	word based	一方、4.2倍を超えると耐加水分解性に <b>優れる</b> が、基材層(B)の凝集強度が低下するため好ましくない。

Table 6: JPCzh-ja task's examples having BLEU and HUMAN scores contradiction (under-translation in word based: example 1 and 2; different translation: example 3 and 4)

1	source	이러한 경우, 수신기는 S707단계로 진행하여 2차 장치가 <b>연결</b> 되었는지 판단한다.
	reference	この場合、受信機は、S707段階において、2次装置が <b>接続</b> しているか否かを判断する。
	char. based	このような場合、受信機は、S707段階に進行して二次装置が <b>連結</b> されたか否かを判断する。
	word based	このような場合、受信機は、S707段階に進行して二次装置が <b>接続</b> されたか否かを判断する。
2	source	[B] 중합성 화합물은, 디펜타에리트릴펜타아크릴레이트와 디펜타에리트릴헥사아크릴레이트의 혼합물인 것이 바람직하다.
	reference	[B] 重合性化合物は、 <b>ジペンタエリスリトールペンタアクリレートとジペンタエリスリトールヘキサアクリレート</b> との混合物であることが好ましい。
	char. based	[B] 重合性化合物は、 <b>ディペンタエリトールペンタアクリレートとディペンターにリトールヘキサアクリレート</b> の混合物であることが好ましい。
	word based	[B] 重合性化合物は、 <b>ペンタエ리트リトールとペンタエ리트リトール</b> ヘキサレートの混合物であることが好ましい。
3	source	이렇게 하면 <b>브레이크</b> 한 부분에서 머더 기판을 단위 기판으로 분단할 수 있다.
	reference	こうすれば <b>ブレイク</b> した部分でマザー基板を単位基板に分断することができる。
	char. based	こうすれば <b>ブレーキ</b> した部分でマザー基板を単位基板に分断することができる。
	word based	こうすれば <b>ブレイク</b> した部分でマザー基板を単位基板に分断することができる。

Table 7: JPCko-ja task's examples having BLEU and HUMAN scores contradiction (different translation: example 1 and 3; un-translation in word based: example 2)

## 4 Conclusion

System descriptions, experimental settings and experimental results of the EHR team are described. We participate in the 3 tasks and submitted 10 systems' outputs. We can observe our re-ranking technique is effective to remove under-translation and over-translation which are in NMT outputs sometimes.

## References

- Hiroshi Echizen-ya and Kenji Araki. 2007. Automatic Evaluation of Machine Translation based on Recursive Acquisition of an Intuitive Common Parts Continuum, *Proceedings of the Eleventh Machine Translation Summit (MT SUMMIT XI)*, Page.151-158.
- Terumasa Ehara. 2015. System Combination of RBMT plus SPE and Preordering plus SMT. *Proceedings of the 2nd Workshop on Asian Translation (WAT2015)*, pages 29–34.
- Terumasa Ehara. 2016. Translation systems and experimental results of the EHR group for WAT2016 tasks. *Proceedings of the 3rd Workshop on Asian Translation (WAT2016)*, pages 111-118.
- Philipp Koehn, Franz J. Och and Daniel Marcu. 2003. Statistical Phrase-Based Translation. *Proceedings of HLTNAACL 2003*, pages 48-54.
- Sadao Kurohashi, Toshihisa Nakamura, Yuji Matsumoto and Makoto Nagao. 1994. Improvements of Japanese morphological analyzer JUMAN. *Proceedings of The International Workshop on Shareable Natural Language Resources*, pages 22-28.
- Minh-Thang Luong, Hieu Pham and Christopher D. Manning. 2015. Effective Approaches to Attention-based Neural Machine Translation, *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1412–1421.
- Toshiaki Nakazawa, Chenchen Ding, Hideya Mino, Isao Goto, Graham Neubig and Sadao Kurohashi. 2016. Overview of the 3rd Workshop on Asian Translation. *Proceedings of the 3rd Workshop on Asian Translation (WAT2016)*, pages 1-46.
- Toshiaki Nakazawa, Shohei Higashiyama, Chenchen Ding, Hideya Mino, Isao Goto, Graham Neubig, Hideto Kazawa, Yusuke Oda, Jun Harashima and Sadao Kurohashi. 2017. Overview of the 4th Workshop on Asian Translation. *Proceedings of the 4th Workshop on Asian Translation (WAT2017)*, pages ??-??.