

PENERJEMAHAN DOKUMEN INGGRIS-INDONESIA MENGGUNAKAN MESIN PENERJEMAH STATISTIK DENGAN *WORD REORDERING* DAN *PHRASE REORDERING*

Hansel Tanuwijaya dan Hisar Maruli Manurung

Fakultas Ilmu Komputer, Universitas Indonesia, Depok, Indonesia
hansel.tanuwijaya@yahoo.com; maruli@cs.ui.ac.id;

Abstrak

Mesin penerjemah merupakan alat penerjemah otomatis pada sebuah teks dari satu bahasa ke bahasa lainnya. Mesin penerjemah statistik adalah sebuah pendekatan mesin penerjemah dengan hasil terjemahan dihasilkan atas dasar model statistik yang parameter-parameternya diambil dari hasil analisis korpus teks *bilingual* (atau paralel). Kualitas hasil terjemahan Bahasa Inggris – Bahasa Indonesia tersebut masih jauh dari sempurna dan memiliki nilai akurasi yang rendah. Diawali dari permasalahan ini, muncullah sebuah ide untuk membuat aturan-aturan restrukturisasi teks pada Bahasa Inggris sesuai dengan struktur Bahasa Indonesia dengan tujuan untuk meningkatkan kualitas dan nilai akurasi hasil terjemahan mesin penerjemah statistik. Aturan restrukturisasi teks tersebut bisa berupa *word reordering*, *phrase reordering*, ataupun keduanya. Dalam penelitian ini, penulis merancang 7 buah aturan *word reordering*, 7 buah aturan *phrase reordering*, dan 2 buah aturan gabungan *phrase reordering* dan *word reordering*. Hasil eksperimen menunjukkan peningkatan akurasi dan kualitas penerjemahan yang efektif diperoleh dengan *word reordering*. *Word reordering* dapat memberikan peningkatan nilai BLEU sebesar 1.3896% (dari 0.1871 menjadi 0.1897) dan nilai NIST sebesar 0.6218% (dari 5.3876 menjadi 5.4211). Pada korpus *bible*, rata-rata nilai peningkatan nilai BLEU yang diperoleh dengan restrukturisasi teks adalah 0.5871% dan untuk nilai NIST terjadi penurunan sebesar 0.0144%. Pada korpus novel, rata-rata nilai peningkatan nilai BLEU yang diperoleh dengan restrukturisasi teks adalah 0.8751% dan untuk nilai NIST terjadi peningkatan sebesar 0.3170%. Besarnya peningkatan dan penurunan yang terjadi pada penelitian ini cenderung kecil (masih di bawah 1%). Hal ini disebabkan aturan penerjemahan Bahasa Inggris-Indonesia menggunakan aturan MD-DM yang melibatkan penukaran kata yang jaraknya dekat sudah tercakup dalam *distortion model* pada mesin penerjemah statistik berdasarkan frase.

Kata kunci : *Bahasa Indonesia, Bahasa Inggris, mesin penerjemah statistik, phrase reordering, word reordering*

1. Pendahuluan

Bahasa merupakan alat komunikasi yang digunakan oleh seseorang untuk menyampaikan ide, gagasan, pengalaman, dan sebagainya kepada orang lain. Bahasa memiliki peranan yang besar dalam kehidupan manusia. Terdapat berbagai ragam macam bahasa yang ada di dunia, salah satunya adalah Bahasa Inggris dan Bahasa Indonesia. Masyarakat Indonesia sekarang ini masih tergolong lemah dalam penguasaan Bahasa Inggris sehingga faktor ini menjadi faktor yang sangat penting dalam terhambatnya penyebaran informasi. Banyak orang seperti: ilmuwan, insinyur, pengusaha harus membaca banyak dokumen dan berkomunikasi dalam Bahasa Inggris. Dengan adanya kemajuan teknologi yang sudah merambah semua bidang, saat ini sedang dikembangkan mesin penerjemah untuk

mengatasi kendala penerjemahan bahasa tersebut. Akan tetapi, kualitas hasil terjemahan tersebut masih jauh dari sempurna dan memiliki nilai akurasi yang rendah. Diawali dari permasalahan ini, muncullah sebuah ide untuk melakukan percobaan agar dapat meningkatkan kualitas dan nilai akurasi hasil terjemahan dari mesin penerjemah statistik dengan menggunakan teknik restrukturisasi teks Bahasa Indonesia.

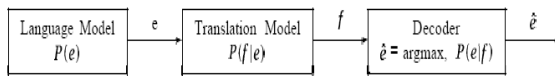
2. Mesin Penerjemah Statistik

Statistical machine translation atau mesin penerjemah statistik merupakan salah satu jenis mesin penerjemah dengan menggunakan pendekatan statistik. Pendekatan statistik yang digunakan adalah konsep probabilitas. Setiap pasangan kalimat (S, T) akan diberikan sebuah $P(T|S)$ yang diinterpretasikan

sebagai distribusi probabilitas yang sebuah penerjemah akan menghasilkan T dalam bahasa tujuan ketika diberikan S dalam bahasa sumber.

2.1 Komponen Mesin Penerjemah Statistik

Menurut Christopher D Manning dan Hinrich Schutze, dalam *statistical machine translation* terdapat tiga buah komponen yang terlibat dalam proses penerjemahan kalimat dari suatu bahasa ke bahasa lain, yaitu: *language model*, *translation model*, dan *decoder* [1].



Gambar 1. Komponen Statistical Machine Translation

Language model digunakan pada aplikasi *Natural Language Processing* seperti *speech recognition*, *part-of-speech tagging* dan *syntactic parsing*.

Language model statistik menetapkan probabilitas $P(w_{1,n})$ ke serangkaian n kata dengan *means* sebuah distribusi probabilitas. Rangkaian-rangkaian tersebut bisa berupa frase-frase atau kalimat-kalimat dan probabilitasnya dapat diperkirakan dari korpus dokumen-dokumen yang besar. Salah satu contoh pendekatan *language model* adalah *n-gram model*.

Probabilitas bersyarat (*conditional*) dapat dihitung dari perhitungan frekuensi *n-gram*:

$$P(w_i | w_{i-(n-1)}, \dots, w_{i-1}) = \frac{\text{count}(w_{i-(n-1)}, w_{i-1}, \dots, w_i)}{\text{count}(w_{i-(n-1)}, \dots, w_{i-1})} \quad (1)$$

Beberapa contoh model bahasa *n-gram* adalah:

- Unigram (1-gram) :
$$P(w_{1,n}) = P(w_1)P(w_2) \dots P(w_n) \quad (2)$$

- Bigram (2-gram) :
$$P(w_{1,n}) = P(w_1)P(w_2 | w_1) \dots P(w_n | w_{n-1}) \quad (3)$$

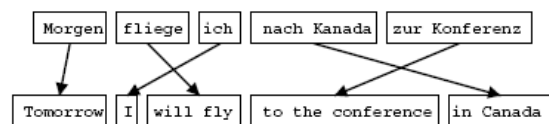
- Trigram (3-gram) :
$$P(w_{1,n}) = P(w_1)P(w_2 | w_1)P(w_3 | w_{1,2}) \dots P(w_n | w_{n-2,n-1}) \quad (4)$$

Language model $P(e)$ berperan untuk menentukan apakah kalimat hasil terjemahan e adalah kalimat Bahasa Inggris yang natural dan memiliki aturan gramatikal yang baik. Semakin tinggi nilai probabilitas dari $P(e)$ menunjukkan bahwa kalimat hasil terjemahan merupakan kalimat

yang dibentuk dengan baik.

Translation model merupakan salah satu komponen penting pada *statistical machine translation* dalam proses penerjemahan.

Proses penerjemahan dalam *phrase-based translation model* dapat dipecah menjadi beberapa bagian yaitu: membagi kalimat bahasa asal menjadi barisan frase, menerjemahkan setiap frase ke bahasa tujuan, dan *reordering* [2].



Gambar 2. Ilustrasi Phrase-Based Translation Model

Komponen terakhir dari mesin penerjemah statistik adalah *decoder*. Fungsi dari *decoder* adalah untuk mencari teks dalam bahasa tujuan (*target language*) yang memiliki probabilitas paling besar dengan pertimbangan faktor *translation model* dan *language model*. Penghitungan \hat{e} dapat dituliskan sebagai berikut:

$$\hat{e} = \arg_e \max P(e|f) = \arg_e \max \frac{P(f|e).P(e)}{P(f)} = \arg_e \max P(f|e)P(e) \quad (5)$$

$\arg_e \max$ mencari e yang dapat memberikan nilai probabilitas terbesar yang diperoleh. Simbol $P(f)$ dapat dihilangkan karena f tetap [1].

2.2 Evaluasi Otomatis

Akurasi dan kualitas dari sebuah *statistical machine translation* pada umumnya dinilai dari hasil terjemahan yang dihasilkan. Penilaian tersebut bisa dilakukan secara manual dan otomatis. Penilaian secara manual tentunya merupakan cara penilaian yang terbaik karena memberikan nilai akurasi yang lebih tepat. Namun penilaian secara manual merupakan cara penilaian yang lambat, mahal, dan lama sebab membutuhkan sumber daya manusia yang banyak. Oleh karena itu, dibutuhkan suatu sistem evaluasi otomatis untuk menentukan kualitas terjemahan yang cepat, murah, dan cukup bagus. Penelitian ini menggunakan evaluasi otomatis yang bernama BLEU dan NIST. BLEU (*Bilingual Evaluation Understudy*) mengukur *modified n-gram precision score* antara hasil terjemahan otomatis dengan terjemahan rujukan dan menggunakan konstanta yang dinamakan *brevery penalty*.

$$BP_{BLEU} = \begin{cases} 1 & \text{if } c > r \\ e^{(1-r/c)} & \text{if } c \leq r \end{cases} \quad (6)$$

$$p_n = \frac{\sum_{C \in \text{corpus}} \sum_{n\text{-gram} \in C} \text{count}_{\text{clip}}(n\text{-gram})}{\sum_{C \in \text{corpus}} \sum_{n\text{-gram} \in C} \text{count}(n\text{-gram})} \quad (7)$$

$$\text{BLEU} = \text{BP}_{\text{BLEU}} \cdot e^{\sum_{n=1}^N w_n \log p_n} \quad (8)$$

Simbol BP merupakan *brevity penalty*, c merupakan jumlah kata dari hasil terjemahan otomatis, r merupakan jumlah kata dari rujukan, dan p_n merupakan *modified precision score*. Nilai w_n adalah 1/N. Standar nilai N untuk BLEU adalah 4, karena nilai presisi BLEU pada umumnya dihitung sampai 4-gram saja. Simbol p_n diperoleh dari jumlah n-gram pada hasil terjemahan yang cocok dengan rujukan dibagi dengan jumlah n-gram pada hasil terjemahan.

NIST (*National Institute of Standards and Technology*) merupakan perkembangan dari BLEU. NIST juga menggunakan penilaian berdasarkan presisi n-grams. NIST menghitung presisi dari n-gram dengan bobot yang berbeda untuk setiap kata. Semakin langka kemunculan suatu kata n-gram maka semakin besar bobot yang diberikan. Nilai NIST didapatkan dari perkalian *brevity penalty* dengan jumlah dari *precision score*. Berbeda dengan BLEU yang memberikan bobot presisi n-grams yang sama [3].

$$\text{BP}_{\text{NIST}} = e^{\left\{ \beta \log^2 \left[\min \left(\frac{L_{\text{sys}}}{L_{\text{ref},1}} \right) \right] \right\}} \quad (9)$$

$$\text{Info}(w_1 \dots w_n) = \log_2 \left(\frac{\text{the \# of occurrences of } w_1 \dots w_n \text{ in all reference}}{\text{the \# of occurrences of } w_1 \dots w_n \text{ in all reference}} \right) \quad (10)$$

$$\text{NIST} = \sum_{n=1}^N \left\{ \frac{\sum \text{Info}(w_1 \dots w_n)}{\text{all } w_1 \dots w_n \text{ that co-occur}} \right\} \cdot \text{BP}_{\text{NIST}} \quad (11)$$

Simbol L_{sys} merupakan jumlah kata yang ada pada hasil terjemahan, L_{ref} merupakan rata – rata dari jumlah kata yang ada pada semua rujukan. Standar nilai N untuk NIST adalah 5, karena nilai presisi NIST pada umumnya dihitung sampai 5-gram saja.

2.3 Word Reordering dan Phrase Reordering

Preprocessing merupakan langkah penting dalam aplikasi *natural language*. Pengaturan ulang urutan kata dan frase sebagai salah satu langkah yang diterapkan pada tahap *preprocessing* dapat menunjukkan peningkatan menurut *automatic*

evaluation pada mesin penerjemah statistik. Hasil dari *preprocessing* digunakan dalam tahap *training* (pembentukan *language model* dan *translation model*) dan tahap *testing* (*decoding* kalimat).

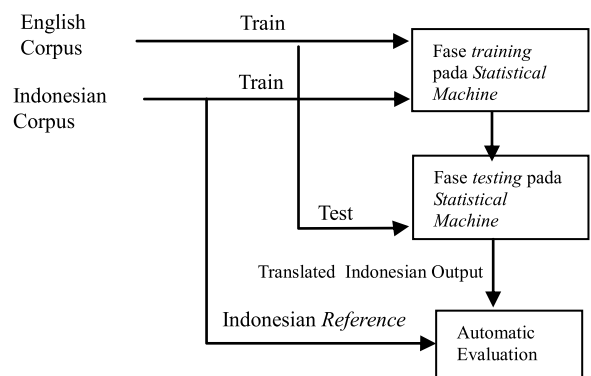
Secara umum, *preprocessing* ini terdiri dari dua tahap, yaitu membagi kalimat yang akan diterjemahkan menjadi barisan kata dan melakukan *word reordering* atau *phrase reordering* pada kalimat bahasa asal (*source language*) berdasarkan struktur kalimat bahasa tujuan (*target language*) [4].

3. Arsitektur Perancangan dan Implementasi Sistem

Pada penelitian ini, digunakan empat buah arsitektur, yaitu: arsitektur *baseline*, arsitektur sistem POS *tagging*, arsitektur sistem *parsing*, dan arsitektur *mixed*.

3.1 Arsitektur Baseline

Pada arsitektur *baseline*, data set korpus paralel Bahasa Inggris dan Bahasa Indonesia dibagi dengan menggunakan metode *K-Fold Cross Validation*. Setelah itu dilakukan proses pembuatan *language model* dan *translation model* dengan data set korpus paralel Bahasa Inggris dan Bahasa Indonesia dalam mesin penerjemah statistik. Kemudian *decoder* dari mesin penerjemah statistik melakukan proses penerjemahan pada data set korpus paralel Bahasa Inggris. Lalu hasil penerjemahan tersebut yang tertulis dalam Bahasa Indonesia dibandingkan dengan data set korpus paralel Bahasa Indonesia yang digunakan sebagai rujukan pada *Automatic Evaluation*.

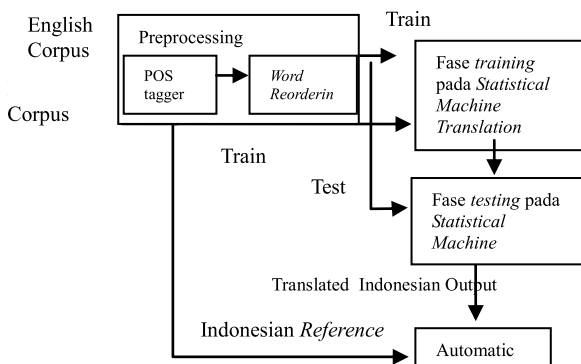


Gambar 3. Arsitektur Baseline

3.2 Arsitektur Sistem POS Tagging

Pada arsitektur sistem POS *tagging*, dilakukan *preprocessing* terlebih dahulu pada data set korpus paralel Bahasa Inggris. Korpus paralel Bahasa Inggris dimasukkan ke dalam POS *tagger* untuk

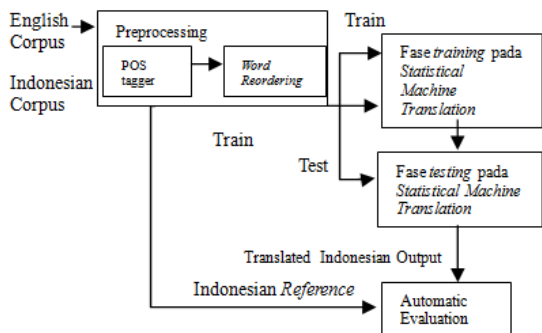
mendapatkan tanda kelas kata pada setiap kata yang terdapat dalam kalimat. Setelah mendapatkan informasi tanda kelas kata, dilakukan proses *word reordering* untuk menukar posisi kata dalam kalimat tersebut. Terdapat tujuh jenis proses *word reordering* yang dilakukan. Setelah dilakukan proses *word reordering*, simbol POS tag yang terdapat pada korpus dihapus dan kemudian melalui tahapan berikutnya yang sama dengan arsitektur *baseline*.



Gambar 4. Arsitektur Sistem POS Tagging

3.3 Arsitektur Sistem Parsing

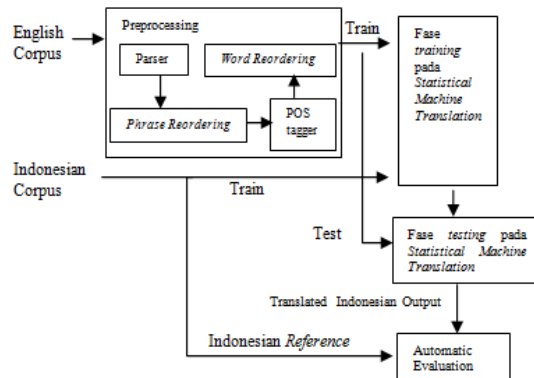
Pada arsitektur sistem *parsing*, dilakukan *preprocessing* terlebih dahulu pada data set korpus paralel Bahasa Inggris. Korpus paralel Bahasa Inggris dimasukkan ke dalam *parser* untuk mendapatkan informasi frase-frase yang terdapat dalam kalimat. Setelah itu, dilakukan beberapa proses *phrase reordering* untuk menukar posisi frase dalam kalimat tersebut. Terdapat tujuh jenis proses *phrase reordering* yang dilakukan. Kemudian data set korpus paralel Bahasa Inggris yang telah melalui *preprocessing* dan Bahasa Indonesia tersebut melalui tahapan berikutnya yang sama dengan arsitektur *baseline*.



Gambar 5. Arsitektur Sistem Parsing

3.4 Arsitektur Mixed

Pada arsitektur *Mixed*, dilakukan *preprocessing* terlebih dahulu pada data set korpus paralel Bahasa Inggris. Korpus paralel Bahasa Inggris dimasukkan ke dalam *parser* untuk mendapatkan informasi frase-frase yang terdapat dalam kalimat. Setelah itu, dilakukan beberapa proses *phrase reordering* untuk menukar posisi frase dalam kalimat tersebut. Proses *phrase reordering* yang diterapkan adalah ADJP-NNx. Setelah proses *phrase reordering*, korpus paralel Bahasa Inggris diproses dengan POS tagger dan dilakukan proses *word reordering*. Terdapat dua macam proses *word reordering* yang diterapkan, yaitu: JJx-NNx dan PRP\$-NNx. Setelah dilakukan proses *word reordering*, simbol POS tag yang terdapat pada korpus dihapus. Kemudian data set korpus paralel Bahasa Inggris yang telah melalui *preprocessing* dan Bahasa Indonesia tersebut melalui tahapan berikutnya yang sama dengan arsitektur *baseline*.



Gambar 6. Arsitektur Mixed

3.5 Korpus Paralel

Korpus paralel adalah dua buah kumpulan dokumen yang memiliki isi yang sama dan ditulis dalam bahasa yang berbeda. Terdapat dua buah korpus paralel Bahasa Inggris dan Bahasa Indonesia yang digunakan dalam penelitian ini, yaitu: korpus paralel *bible* dan korpus paralel novel. Sebelum kedua korpus paralel tersebut digunakan dalam penelitian, terlebih dahulu dilakukan pengecekan tanda baca dan karakter pada korpus paralel. Masing – masing korpus paralel tersebut memiliki 12000 baris. Kemudian 12000 baris dibagi menjadi 6 buah *fold* dimana masing – masing *fold* berjumlah 2000 baris. Penelitian ini menggunakan metode *6-fold cross validation* di mana setiap percobaannya melibatkan 1 *fold testing* dan 5 *fold training*.

3.6 Rancangan Word Reordering dan Phrase Reordering

Penerjemahan kalimat Bahasa Inggris ke Bahasa Indonesia akan melibatkan aturan tata bahasa tertentu. Salah satu aturan tata bahasa yang penting adalah aturan MD-DM [5]. M adalah menerangkan dan D adalah diterangkan. Sebuah frase pada Bahasa Inggris dibentuk dengan aturan menerangkan-diterangkan (MD), sedangkan sebuah frase pada Bahasa Indonesia frase dibentuk dengan aturan diterangkan-menerangkan (DM). Pola MD berarti kata yang berada di depan menerangkan kata yang berada di belakangnya sedangkan pola DM berarti kata yang berada di belakang menerangkan kata yang berada di depannya. Kata yang diterangkan menjadi unsur utama frase dan kata yang menerangkan menjadi unsur penjelas, sehingga penerjemahan kata dari Bahasa Inggris menjadi Bahasa Indonesia akan mengikuti aturan MD-DM. Contoh: *intelligent student* (MD) menjadi siswa pintar (DM), *branch office* (MD) menjadi kantor cabang (DM).

Penulis merancang 3 buah aturan pokok dalam proses *word reordering*, yaitu:

a. JJx-NNx *reordering*

JJx-NNx *reordering* adalah penukaran urutan kata sifat dengan kata benda. Kata sifat dalam Bahasa Inggris ditandai dengan tag /JJ, /JJN, /JJS dan /JJS. Penulis menyebut kelompok *tag* untuk kata sifat tersebut dengan JJx. Kata benda dalam Bahasa Inggris ditandai dengan tag /NN, /NNP, /NNS dan /NNPS. Penulis menyebut kelompok *tag* untuk kata benda tersebut dengan NNx. Apabila pada suatu kalimat terdapat kata sifat yang diikuti dengan kata benda, posisi kedua kata tersebut ditukar.

b. PRP\$-NNx *reordering*

PRP\$-NNx *reordering* adalah penukaran urutan kata kepemilikan dengan kata benda. Kata kepemilikan dalam Bahasa Inggris ditandai dengan tag /PRP\$. Kata benda dalam Bahasa Inggris ditandai dengan tag /NN, /NNP, /NNS dan /NNPS. Penulis menyebut kelompok *tag* untuk kata benda tersebut dengan NNx. Apabila pada suatu kalimat terdapat kata kepemilikan yang diikuti dengan kata benda, posisi kedua kata tersebut ditukar.

c. NNx-NNx *reordering*

NNx-NNx *reordering* adalah penukaran urutan kata benda dengan kata benda. Kata benda dalam Bahasa Inggris ditandai dengan tag /NN, /NNP, /NNS dan /NNPS. Penulis menyebut kelompok *tag* untuk kata benda tersebut dengan NNx. Apabila pada suatu kalimat terdapat kata benda yang diikuti dengan kata benda, posisi kedua kata tersebut ditukar.

Penulis mendapatkan 3 aturan tersebut dari studi literatur [5], analisis empiris korpus paralel secara umum, dan melalui diskusi dengan pembimbing. Pada tahap implementasi, eksperimen *word reordering* dilakukan dengan mencoba semua kemungkinan dari 3 buah variasi *word reordering* tersebut. Oleh karena itu terdapat 7 buah kombinasi aturan *word reordering*, yaitu:

- JJx-NNx *reordering*
- NNx-NNx *reordering*
- PRP\$-NNx *reordering*
- JJx-NNx dan NNx-NNx *reordering*
- JJx-NNx dan PRP\$-NNx *reordering*
- NNx-NNx dan PRP\$-NNx *reordering*
- JJx-NNx, NNx-NNx dan PRP\$-NNx *reordering*

Phrase reordering adalah penukaran ulang susunan frase atau kelompok kata dalam sebuah kalimat. Dalam penelitian ini, penukaran ulang susunan frase dalam kalimat Bahasa Inggris dilakukan agar kalimat tersebut dapat mempunyai susunan kalimat yang semirip mungkin dengan aturan Bahasa Indonesia.

Penulis merancang 3 buah aturan pokok dalam proses *phrase reordering*, yaitu:

a. ADJP-NP *reordering*

ADJP-NP *reordering* adalah penukaran urutan frase sifat atau frase adjektival dengan frase benda. ADJP merupakan simbol untuk frase sifat dan NP merupakan simbol untuk frase benda. Apabila pada suatu kalimat terdapat frase sifat yang diikuti dengan frase benda, posisi kedua frase tersebut ditukar.

b. NP-NP *reordering*

NP-NP *reordering* adalah penukaran urutan frase benda dengan frase benda. NP merupakan simbol untuk frase benda. Apabila pada suatu kalimat terdapat frase benda yang diikuti dengan frase benda, posisi kedua frase tersebut ditukar.

c. ADJP-NNx *reordering*

ADJP-NNx *reordering* adalah penukaran urutan frase sifat dengan kata benda. ADJP merupakan simbol untuk frase sifat dan NNx merupakan simbol untuk kata benda. Apabila pada suatu kalimat terdapat frase sifat yang diikuti dengan kata benda, posisi frase sifat dan kata benda tersebut ditukar.

Penulis mendapatkan 3 aturan tersebut dari analisis empiris korpus secara umum. Pada tahap implementasi, eksperimen *phrase reordering* dilakukan dengan mencoba semua kemungkinan dari 3 buah variasi *phrase reordering* tersebut. Sehingga terdapat 7 buah kombinasi aturan *phrase reordering*, yaitu:

- ADJP-NNx *reordering*
- NP-NP *reordering*

- ADJP-NP reordering
- ADJP-NNx dan NP-NP reordering
- ADJP-NNx dan ADJP-NP reordering
- ADJP-NP dan NP-NP reordering
- ADJP-NNx, NP-NP reordering, dan ADJP-NP reordering

3.7 Tools yang Digunakan

Penelitian ini menggunakan beberapa tools seperti:

- POS Tagger: Stanford POS Tagger [6].
- Parser: Stanford Parser [7].
- Statistical Machine Translation : MOSES, SRILM, dan GIZA++ [8].
- Automatic Evaluation: Mteval-v11b.pl [9].

Selain tools di atas terdapat tools lain yang dibuat dengan menggunakan bahasa pemrograman Perl dan Java untuk persiapan awal dokumen, cleaning dokumen, word reordering, phrase reordering, dan format SGM untuk evaluasi.

4. Hasil Penelitian

Berikut merupakan rangkuman hasil penelitian arsitektur baseline, POS tagging, parsing, dan mixed.

Tabel 1. Rangkuman Hasil Eksperimen Nilai BLEU dan NIST pada Penelitian

BLEU Inggris - Indonesia	Nilai	Perubahan (%)	NIST Inggris-Indonesia	Nilai	Perubahan
Baseline	0.1871		Baseline	5.3876	
JJx-NNx	0.1890	+1.0155 %	JJx-NNx	5.4108	+0.4306 %
JJx-NNx dan NNx-NNx	0.1897	+1.3896 %	JJx-NNx dan NNx-NNx	5.4211	+0.6218 %
JJx-NNx dan PRPS-NNx	0.1892	+1.1224 %	JJx-NNx dan PRPS-NNx	5.4166	+0.5383 %
JJx-NNx, NNx-NNx, dan PRPS-NNx	0.1906	+1.870 %	JJx-NNx, NNx-NNx, dan PRPS-NNx	5.4317	+0.8185 %
ADJP-NNx	0.1873	+0.1069 %	ADJP-NNx	5.3741	-0.2506%
ADJP-NP	0.1872	+0.0534 %	ADJP-NP	5.3741	-0.2506%
ADJP-NNx dan JJx-NNx	0.1891	+1.0689 %	ADJP-NNx dan JJx-NNx	5.4081	0.3796%
ADJP-NNx dan PRPS-NNx	0.1878	+0.3741 %	ADJP-NNx dan PRPS-NNx	5.401	+0.2487 %

Kotak yang diarsir hitam menandakan nilai yang lebih tinggi dari nilai baseline. Kotak yang diarsir

abu-abu menandakan nilai yang lebih rendah dari nilai baseline.

Pada korpus bible, rata-rata nilai peningkatan nilai BLEU yang diperoleh adalah 0.5871% dan untuk nilai NIST terjadi penurunan sebesar 0.0144%. Pada korpus novel, rata-rata nilai peningkatan nilai BLEU yang diperoleh adalah 0.8751% dan untuk nilai NIST terjadi peningkatan sebesar 0.3170%.

Rata-rata jumlah peningkatan dan penurunan yang terjadi pada penelitian ini, jika dibandingkan dengan percobaan lain yang pernah dilakukan, adalah cenderung kecil (masih di bawah 1%). Hal ini dikarenakan aturan penerjemahan Bahasa Inggris-Indonesia menggunakan aturan MD-DM yang melibatkan penukaran kata yang jaraknya dekat (bersifat local reordering). Penukaran kata tersebut mirip dengan reordering yang sudah dimiliki oleh mesin penerjemah statistik sehingga proses reordering Bahasa Inggris-Indonesia dengan menggunakan aturan MD-DM tidak terlalu berpengaruh. Berbeda dengan aturan reordering pada bahasa lain (contoh: Jerman - Inggris, Inggris - India) yang melibatkan penukaran kata yang jauh, penukaran kata yang jauh tersebut tidak bisa ditangani oleh reordering yang ada pada mesin penerjemah statistik. Oleh karena itu, aturan reordering pada bahasa lain tersebut memberikan dampak perubahan nilai BLEU dan NIST yang lebih besar.

Penelitian Sangodkar pada Bahasa Inggris - India. Sangodkar menggunakan korpus EILMT dan IIT serta standar penilaian BLEU dan NIST.

Tabel 2. Hasil Eksperimen Nilai BLEU dan NIST Bahasa Inggris - India

Bleu	EILMT		IIT	
	Baseline	Reorder Verb	Baseline	Reorder Verb
Data Set 1	0.1488	0.1751	0.0815	0.0836
Perubahan (%)		17.67%		2.58%
Data Set 2	0.145	0.1601	0.0842	0.0853
Perubahan		9.43%		1.30%
NIST	Baseline	Reorder Verb	Baseline	Reorder Verb
Data Set 1	4.76	4.8539	3.9036	3.7335
Perubahan (%)		1.97%		-4.36%
Data Set 2	4.7287	4.6923	4.2426	4.014
Perubahan (%)		-0.77%		-5.39%

Dari gambar di atas, percobaan hasil eksperimen Sangodkar menunjukkan adanya peningkatan pada nilai BLEU. Akan tetapi pada percobaan tersebut terdapat penurunan nilai NIST pada korpus EILMT pada data set 2 serta korpus IIT pada data set 1 dan data set 2.

Rata-rata peningkatan nilai BLEU yang diperoleh pada penerjemahan Bahasa Inggris - India yang diperoleh dari korpus EILMT adalah 13.5524%. dan pada korpus IIT adalah 1.9383%. Rata-rata peningkatan nilai NIST yang diperoleh pada penerjemahan Bahasa Inggris - India yang diperoleh dari korpus EILMT adalah 0.3007%. sedangkan untuk korpus IIT terjadi penurunan 4.87285% [10].

Kemiripan hasil yang diperoleh pada penelitian ini dengan hasil yang didapatkan dari penelitian Sangodkar pada Bahasa Inggris – India adalah terjadi peningkatan untuk nilai BLEU dan terdapat variasi peningkatan dan penurunan nilai NIST pada eksperimen.

5. Kesimpulan dan Saran

Kesimpulan dari penelitian yang telah dilakukan penulis adalah sebagai berikut:

- Pada penelitian ini telah dirancang dan dilakukan percobaan dengan 7 buah aturan *word reordering*, 7 buah aturan *phrase reordering*, dan 2 buah aturan gabungan *phrase reordering* dan *word reordering*.
- Pada korpus *bible*, rata-rata nilai peningkatan nilai BLEU yang diperoleh rekstrukturisasi teks adalah 0.5871% dan untuk nilai NIST terjadi penurunan sebesar 0.0144%.
- Pada korpus novel, rata-rata nilai peningkatan nilai BLEU yang diperoleh dari rekstrukturisasi teks adalah 0.8751% dan untuk nilai NIST terjadi peningkatan sebesar 0.3170%.
- Pada arsitektur sistem POS *tagging*, terdapat tiga buah aturan *word reordering* yang dapat memberikan peningkatan kualitas dan akurasi mesin penerjemah statistik, yaitu: penukaran kata sifat dengan kata benda (JJx-NNx), penukaran kata sifat dengan kata benda dan kata benda dengan kata benda (JJx-NNx dan NNx-NNx), penukaran kata sifat dengan kata benda dan kata kepunyaan dengan kata benda (JJx-NNx dan PRPS-NNx). Ketiga aturan tersebut menunjukkan adanya peningkatan baik berdasarkan nilai BLEU maupun berdasarkan nilai NIST pada korpus *bible* dan novel.
- Pada arsitektur sistem *parsing*, dari semua rancangan *phrase reordering* yang dilakukan tidak ada aturan *phrase reordering* yang dapat memberikan peningkatan kualitas dan akurasi mesin penerjemah statistik baik berdasarkan nilai BLEU dan nilai NIST pada korpus *bible* dan novel.
- Pada arsitektur sistem *mixed*, dari semua rancangan *phrase and word reordering* yang

dilakukan tidak ada aturan *phrase and word reordering* yang dapat memberikan peningkatan kualitas dan akurasi mesin penerjemah statistik baik berdasarkan nilai BLEU dan nilai NIST pada korpus *bible* dan novel.

- Rata-rata jumlah peningkatan dan penurunan yang terjadi pada penelitian ini jika dibandingkan dengan percobaan lain yang pernah dilakukan adalah cenderung kecil (masih di bawah 1%). Hal ini dikarenakan aturan penerjemahan Bahasa Inggris-Indonesia menggunakan aturan MD-DM yang melibatkan penukaran kata yang jaraknya dekat sudah tercakup dalam *distortion model* pada mesin penerjemah statistik berdasarkan frase.

Saran yang diberikan penulis untuk pengembangan selanjutnya dari penelitian ini adalah:

- Penelitian lebih lanjut dengan menggunakan mesin penerjemah statistik dan *tools* lain.
- Penelitian lebih lanjut dalam pendefinisian aturan *word reordering* dan *phrase reordering* sangat disarankan karena masih terdapat aturan restrukturisasi ataupun metode lain yang mungkin belum diteliti dan mungkin dapat memberikan peningkatan nilai akurasi yang lebih tinggi dari yang diberikan laporan ini.
- Penelitian lebih lanjut dengan menggunakan standar penilaian akurasi dan kualitas penerjemahan selain BLEU dan NIST.
- Penelitian lebih lanjut dengan menggunakan gabungan domain korpus paralel *bible* dan novel.

REFERENSI

- [1] Manning, Christopher D. and Hinrich Schutze. *Foundations of Statistical Machine Translations*. MIT Press: England. 2000.
- [2] Koehn, Phillip. *Statistical Machine Translation: Lecture 3 Word Alignment and Phrase Models*. Edinburgh. 2005.
- [3] Ying Zhang, Stephan Vogel, Alex Waibel. "Interpreting Bleu/NIST scores: how much improvement do we need to have a better system?", *Proceedings of LREC 2004*, Lisbon, Portugal. 2004.
- [4] Zwarts, Simon and Mark Dras. "Syntax based word reordering in phrase-based statistical machine translation: why does it work?", *Proceedings of MT Summit*, pages 559–566, Copenhagen, Denmark. 2007.
- [5] Alwi, Hasan. Et.al. *Tata Bahasa Baku Indonesia*, 3rd ed. Balai Pustaka: Jakarta. 2003.
- [6] Stanford University. *Stanford Log-linear Part-Of-Speech Tagger*. 2008.

- <http://nlp.stanford.edu/software/tagger.shtml>.
- [7] Stanford University. *The Stanford Parser: A statistical parser*. 2008.
<http://nlp.stanford.edu/software/lex-parser.shtml>
- [8] MOSES. 2009.
<http://www.statmt.org/moses/index.html>
- [9] MTEVAL-v11b. 2008.
[ftp://jaguar.ncsl.nist.gov/mt/resources/mteval-](ftp://jaguar.ncsl.nist.gov/mt/resources/mteval-v11b.pl)
- [v11b.pl](ftp://jaguar.ncsl.nist.gov/mt/resources/mteval-v11b.pl)
- [10] Sangodkar, Amit; Vasudevan N., and Om P. Damani. "Statistical machine translation with rule based re-ordering of source sentences", *Proceedings of ICON-2008: 6th International Conference on Natural Language Processing Macmillan Publishers, India*. 2008.