

Abstrak

Mesin translasi berbasis statistik merupakan suatu teknik pemrosesan bahasa alami berdasarkan ilmu statistik dengan menggunakan model pembelajaran dari sekumpulan kalimat sebagai data latihnya (*corpus*). Salah satu faktor yang sangat berpengaruh pada proses pelatihan dan menentukan kualitas hasil translasi adalah proses penerjemahan frasa-frasa yang terdapat pada *corpus*.

Dalam tugas akhir ini dilakukan pengimplementasian aturan koreksi level frasa pada mesin translasi berbasis statistik untuk memahami dan mengetahui seberapa besar pengaruh aturan ini pada kualitas hasil translasi mesin translasi berbasis statistik jika digunakan untuk menerjemahkan kalimat-kalimat dari Bahasa Inggris ke Bahasa Indonesia. Pengimplementasian dilakukan dengan membuat modul aturan koreksi level frasa, mengintegrasikannya pada mesin translasi berbasis statistik, dan kemudian membandingkan kualitas hasil terjemahan mesin translasi berbasis statistik ber-aturan koreksi level frasa tersebut dengan hasil terjemahan beberapa mesin translasi, serta melakukan analisa terhadap hasil perbandingan tersebut.

Pada penelitian ini didapatkan bahwa aturan koreksi level frasa menurunkan kualitas hasil translasi mesin translasi berbasis statistik. Hal ini dikarenakan sifat *context-independent* dan *partial monotone alignment* dari aturan tersebut, dan juga karena sudah adanya penanganan koreksi level frasa yang lebih baik pada mesin translasi berbasis statistik.

Kata kunci: mesin translasi berbasis statistik, corpus, frasa, aturan koreksi level frasa.