

Penerjemahan dokumen inggris-indonesia menggunakan mesin penerjemah statistik dengan word reordering dan phrase reordering

Hansel Tanuwijaya, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20335646&lokasi=lokal>

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.