

Penerapan centering theory dalam analisis koherensi pada hasil peringkasan ekstraktif = Implementation of centering theory for coherence analysis of extractive summarization result / Arlisa Yuliatwati

Arlisa Yuliatwati, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20350689&lokasi=lokal>

Abstrak

ABSTRAK

Peringkasan teks otomatis adalah proses menghasilkan versi sederhana dari sebuah atau sekumpulan dokumen yang membahas topik tertentu. Salah satu teknik peringkasan yang telah banyak diteliti dan cukup mudah diterapkan adalah peringkasan secara ekstraktif. Karena teknik ini melakukan peringkasan dengan mengambil bagian-bagian penting dari dokumen yang diringkas kemudian merangkaikannya kembali sebagai sebuah ringkasan, salah satu permasalahan yang muncul adalah kurangnya keterkaitan atau koherensi antar bagian penyusun ringkasan. Dalam penelitian ini, dilakukan rekonstruksi urutan kalimat hasil peringkasan ekstraktif khususnya pada koleksi dokumen berita. Pengurutan dilakukan dengan menjaga keterkaitan entitas antar pasangan kalimat bertetangga berdasarkan prinsip Continuity yang terdapat pada konsep Centering Theory sembari mempertahankan urutan relatif setiap kalimat dalam ringkasan untuk menjaga alur penyampaian informasi dalam berita. Kedua pertimbangan tersebut digunakan sebagai fungsi fitness pada algoritma genetik yang dipergunakan dalam melakukan rekonstruksi urutan kalimat ringkasan. Berdasarkan hasil evaluasi oleh manusia, penggunaan komposisi bobot sebesar 75% pertimbangan Continuity dan 25% pertimbangan urutan relatif setiap kalimat, mampu menghasilkan urutan kalimat ringkasan yang memiliki kecenderungan dinilai sebagai urutan yang dapat diterima oleh para penilai.

ABSTRACT

Automatic text summarization is a process of producing a simplified version of a document or a set of documents about specific topic. A widely studied and fairly easy to implement is the extractive summarization technique. Because this technique works by choosing the most important parts of the document(s) as a summary, this can lead to the lack of coherence in the summary itself. In this study, principle of Continuity in the concept of Centering Theory is used to maintain the entity coherence between sentences in the summary from extractive news document summarization while maintaining the relative order of each sentence in the summary to keep the flow of information delivery. These considerations are used as a fitness function for genetic algorithm that is used to reconstruct the order of sentences in the summary. Based on the evaluation result by human judgment, the use of 75% principle of Continuity and 25% relative order in each sentence in

the summary, is able to produce an acceptable sentence order.