

Pengembangan lanjut pengurai morfologi untuk bahasa Indonesia berbasis morfologi dua tingkat menggunakan XFST dan LEXC

Pisceldo, Femphy, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=123826&lokasi=lokal>

Abstrak

Pengurai morfologi merupakan suatu program yang digunakan untuk melakukan pengenalan kata dan membagi kata menjadi satu atau lebih morfem dengan memberikan analisis morfologi yang sesuai untuk kata tersebut. Pengurai morfologi yang dikembangkan pada penelitian untuk Tugas Akhir ini adalah pengembangan dari pengurai morfologi yang telah dikembangkan Hendra Hartono pada tahun 2002. Pengurai morfologi ini dikembangkan untuk bahasa Indonesia dengan memanfaatkan prinsip morfologi dua tingkat (two-level morphology) yang lazim dipakai untuk pengurai-pengurai morfologi pada bahasa-bahasa lainnya.

Prinsip morfologi dua tingkat (two-level morphology) ini memanfaatkan finite-state transducers dalam pemodelannya. Morfologi bahasa Indonesia yang dicakup dalam penelitian ini berkisar antara kata dasar, kata imbuhan dari penggabungan kata dasar dengan awalan (prefiks), akhiran (sufiks), maupun gabungan pengimbuhanan (konfiks), hingga kata ulang sejati, kata ulang sebagian, dan kata ulang berimbuhan. Morfologi bahasa Indonesia ini dibuat berdasarkan tata bahasa baku bahasa Indonesia. Perancangan untuk pengurai morfologi ini melingkupi perancangan lexicon, tags, aturan-aturan morfotaktik, hingga aturan-aturan morfofonemik.

Hasil rancangan tersebut kemudian diimplementasikan dengan Xerox Finite-State Tool (XFST) dan Finite-State Lexicon Compiler (LEXC). Setelah tahapan implementasi dilakukan, hasilnya diujicobakan dengan berbagai test cases yang representatif sesuai apa yang dikembangkan dalam penelitian ini. Ujicoba dilakukan terhadap 420 test cases dan hanya 8 test cases yang mengeluarkan hasil yang salah. Kata kunci: Pengurai morfologi, two-level morphology, finite-state transducers, tata bahasa baku bahasa Indonesia, lexicon, tags, morfotaktik, morfofonemik, XFST, dan LEXC.