

## BAB 1 PENDAHULUAN

Pada bab ini, penulis akan membahas beberapa hal yaitu latar belakang dari permasalahan, perumusan masalah, tujuan penelitian, ruang lingkup dan batasan penelitian, metode penelitian, serta sistematika penulisan laporan tugas akhir.

### 1.1 Latar Belakang Masalah

Dalam bidang pemrosesan bahasa alami, diperlukan suatu pengurai yang dapat memberi informasi tentang proses pembentukan suatu kata. Pengurai tersebut dapat memberikan berbagai macam informasi perihal terbentuknya suatu kata dari aspek linguistik. Dengan pengurai itu juga dapat dipelajari struktur dan susunan suatu kata. Pengurai semacam itu dinamakan pengurai morfologi.

Pengurai morfologi dalam penggunaannya dapat dimanfaatkan untuk melihat bagaimana suatu kata terbentuk, apa saja kata dasarnya dan apa saja informasi linguistik yang dikandung kata tersebut. Dengan segala informasi yang dikandung suatu pengurai morfologi, diharapkan dapat membantu dalam mempelajari suatu kata dan proses pembentukannya. Lebih jauh lagi pengurai morfologi dapat dijadikan pijakan dan referensi yang cukup baik untuk pengembangan *stemmer*, *parser*, maupun perangkat-perangkat untuk pemrosesan bahasa alami lainnya.

Untuk bahasa Indonesia, pengembangan pengurai morfologi ini sendiri pernah dikembangkan oleh Hendra Hartono [HART02]. Penelitian tersebut menghasilkan suatu pengurai morfologi dalam bahasa Indonesia yang diimplementasikan dengan menggunakan PC-KIMMO dan menggunakan model morfologi dua tingkat (*two-level morphology*).

Penelitian yang dilakukan kali ini dapat dikatakan sebagai pengembangan dari penelitian Hendra Hartono kala itu. Hanya saja kali ini menggunakan basis yang berbeda untuk mengimplementasikannya dan merancang hasil analisis morfologi yang diusahakan lebih mendalam lagi dari sebelumnya.

## 1.2 Perumusan Masalah

Pengurai morfologi yang ingin dibuat adalah pengurai morfologi untuk bahasa Indonesia menggunakan prinsip morfologi dua tingkat (*two-level morphology*) yang memanfaatkan *finite-state transducers* untuk pemodelannya. Implementasi menggunakan Xerox Finite State Tool (XFST) dan Finite State Lexicon Compiler (LEXC). *Lexicon* untuk penelitian ini didapatkan dari kumpulan kata dasar (*lemma*) WordNet Indonesia yang sudah dikumpulkan dari berbagai sumber kamus termasuk Kamus Besar Bahasa Indonesia (KBBI).

## 1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah menghasilkan suatu pengurai morfologi dalam bahasa Indonesia yang dapat menghasilkan analisis morfologi dari suatu kata. Kata yang dimasukkan dalam lexicon adalah kata dasar, sedangkan kata turunan akan dihasilkan dari kata dasar melalui tahapan pembentukannya.

Pengurai ini diharapkan dapat menghasilkan analisis morfologi yang mendalam dari kumpulan kata-kata yang cukup banyak dalam bahasa Indonesia. Penerapan pengurai morfologi dikembangkan dengan morfologi dua tingkat (*two-level morphology*) dan diimplementasikan dengan XFST dan LEXC.

## 1.4 Ruang Lingkup dan Batasan Penelitian

Penelitian ini dibatasi hanya untuk menguraikan kata-kata dalam bahasa Indonesia. Kata yang dapat dihasilkan dibatasi pada kata-kata yang dapat dibentuk dari kumpulan kata dasar (*lemma*) yang diperoleh dari Kamus Besar Bahasa Indonesia dalam jaringan (KBBI Daring) [KBBD08]. Analisis yang dapat dihasilkan terbatas pada kata dasar, kata imbuhan yang dihasilkan dari penggabungan kata dasar dengan awalan (prefiks), akhiran (sufiks), maupun gabungan imbuhan (konfiks), kemudian juga kata ulang sejati, kata ulang sebagian, dan kata ulang berimbuhan. Perlu diperhatikan analisis tidak mencakup kata majemuk dan kata imbuhan yang merupakan hasil penggabungan kata dasar dengan sisipan (infiks). Hasil analisis yang dihasilkan dibatasi sesuai tags yang dirancang pada subbab 3.3.

## 1.5 Metode Penelitian

Penelitian ini didahului dengan studi pustaka di mana pada tahapan tersebut dipelajari konsep-konsep yang menjadi landasan teori dari penelitian ini. Pembelajaran literatur dan konsep-konsep didapatkan dari berbagai sumber rujukan dalam bidang pemrosesan bahasa alami. Di samping itu, pada tahapan studi pustaka juga dipelajari prinsip-prinsip pembentukan kata dalam bahasa Indonesia dan juga studi kata-kata dalam bahasa Indonesia itu sendiri.

Setelah tahapan studi pustaka dilaksanakan, penelitian ini berlanjut ke tahapan perancangan. Tahapan perancangan mencakup perancangan *lexicon*, *tags*, aturan-aturan morfotaktik, dan aturan-aturan morfofonemik. Tahapan perancangan ini juga mencakup rancangan skema umum pengurai morfologi ini.

Hampir berjalan secara paralel dengan tahapan perancangan juga dilaksanakan tahapan studi *tools* untuk implementasi. Studi ini mencakup pembelajaran *tools* yaitu XFST dan LEXC yang akan digunakan untuk implementasi pengurai morfologi tersebut.

Kemudian barulah penelitian ini memasuki tahapan implementasi secara keseluruhan dengan XFST dan LEXC dan kemudian dilanjutkan dengan pengujian cobaan pengurai morfologi ini masih di dalam XFST dan LEXC.

## 1.6 Sistematika Penulisan

Laporan Tugas Akhir (TA) ini memiliki sistematika penulisan yang secara umum dapat dibagi dalam bab-bab sebagai berikut:

### BAB 1            PENDAHULUAN

Bab ini membahas tentang latar belakang, tujuan penelitian, rumusan masalah, ruang lingkup dan batasan dari penelitian ini, kemudian metode yang digunakan dalam penelitian, dan

juga sistematika dalam penulisa Tugas Akhir (TA) ini.

## BAB 2 LANDASAN TEORI

Bab ini menjelaskan tentang landasan teori yang telah ada sebelumnya dan menjadi pijakan dalam penelitian ini. Termasuk di dalamnya penjelasan tentang morfologi dan pengurai morfologi, serta penjelasan tentang morfologi dua tingkat (*two-level morphology*) dan *finite-state transducers* yang menjadi pijakan teori yang digunakan dalam penelitian ini.

## BAB 3 RANCANGAN PENGURAI MORFOLOGI BAHASA INDONESIA

Bab ini menjelaskan semua perancangan yang dibutuhkan untuk membangun suatu pengurai morfologi dalam bahasa Indonesia. Di dalamnya tercakup segala bentuk perancangan yang dilakukan mulai dari perancangan *lexicon*, *tags*, aturan-aturan morfotaktik, hingga aturan-aturan morfofonemik.

## BAB 4 IMPEMENTASI PENGURAI MORFOLOGI BAHASA INDONESIA

Bab ini berisi penjelasan tentang implementasi yang dilakukan dalam mengembangkan pengurai morfologi dalam bahasa

Indonesia. Termasuk di dalamnya pembahasan tentang XFST dan LEXC yang menjadi basis yang digunakan dalam implementasi pengurai morfologi ini.

## BAB 5 HASIL DAN ANALISIS PENGUJIAN

Bab ini berisi pengujian yang dilakukan terhadap pengurai morfologi bahasa Indonesia yang telah dikembangkan. Di dalamnya juga tercakup analisis yang dilakukan dari hasil pengujian tersebut.

## BAB 6 PENUTUP

Bab ini memaparkan kesimpulan dan saran yang dapat dirumuskan setelah menjalani penelitian ini. Bab ini juga berisi hambatan yang ditemui dalam penelitian ini dan juga *future work* yang bisa dikembangkan dari penelitian ini.