

**PENGEMBANGAN LANJUT PENGURAI MORFOLOGI
UNTUK BAHASA INDONESIA BERBASISKAN MORFOLOGI
DUA TINGKAT MENGGUNAKAN XFST DAN LEXC**

TUGAS AKHIR

**FEMPHY PISCELDO
1204000335**



**UNIVERSITAS INDONESIA
FAKULTAS ILMU KOMPUTER
JURUSAN ILMU KOMPUTER
DEPOK
JULI 2008**

HALAMAN PENGESAHAN

Tugas akhir ini diajukan oleh

Nama : Femphy Pisceldo

NPM : 1204000335

Program Studi : Ilmu Komputer

Judul Tugas Akhir : Pengembangan Lanjut Pengurai Morfologi untuk Bahasa Indonesia Berbasiskan Morfologi Dua Tingkat Menggunakan XFST dan LEXC

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Ilmu Komputer pada Program Studi Ilmu Komputer Fakultas Ilmu Komputer Universitas Indonesia.

DEWAN PENGUJI

Pembimbing : Hisar Maruli Manurung (.....tandatangan.....)

Penguji : Ade Azurat (.....tandatangan.....)

Penguji : Ari Saptawijaya (.....tandatangan.....)

Ditetapkan di : Depok

Tanggal : 24 Juli 2008

KATA PENGANTAR

Pertama-tama penulis ingin memanjatkan puji syukur kepada Allah SWT karena hanya atas berkat rahmat dan karunia-Nya kami dapat menyelesaikan penelitian Tugas Akhir (TA) dan laporannya dengan baik sesuai kaidah yang telah ditentukan.

Penulisan Tugas Akhir ini bertujuan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Indonesia. Penulis menyadari bahwa tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan hingga masa penyusunan laporan Tugas Akhir (TA) ini, sangatlah sulit untuk menyelesaikannya. Oleh karena itu penulis ingin mengucapkan terima kasih kepada:

1. Allah SWT atas berkat rahmat dan karunia yang selalu dilimpahkan.
2. Junjungan besar Nabi Muhammad SAW.
3. Orang tua, sanak saudara, dan keluarga yang selalu tanpa henti memberikan dukungan, semangat, dan motivasi dalam membantu saya melalui kerja praktik ini.
4. Bapak Hisar Maruli Manurung selaku pembimbing Tugas Akhir (TA).
5. Ibu Mirna A. selaku pembimbing LAB Perolehan Informasi.
6. Teman-teman dari LAB Perolehan Informasi.
7. Rahmad Mahendra yang telah banyak membantu perancangan pada penelitian ini.
8. Desmond Darma Putra yang telah membantu dalam mendapatkan daftar kata dasar.
9. Bapak Wayan Arka untuk konsultasinya perihal bahasa Indonesia.
10. Sahabat-sahabat saya Ilham Aji Pratomo, Rakhmat Adhi Pratama, Aditya Nugraha, Hendra Dwi Hadmanto, Edwin Kurniawan, Dani Supriadi, Arief Ristanto, Agung Widiyanto, Gita Lystia, Sari Dwi Handiny, Hadaiq Rolis,

M. Edwin D.P, Franova Herdiyanto, dan sahabat-sahabat saya yang lain yang telah banyak membantu dalam menyelesaikan Tugas Akhir (TA) ini.

11. Teman-teman Fakultas Ilmu Komputer UI angkatan 2004 yang telah banyak memberikan dukungan selama penyelesaian Tugas Akhir (TA) ini.
12. Seluruh dosen, asisten dosen, dan teman-teman yang pernah belajar bersama dalam suatu mata perkuliahan.
13. Segenap *civitas academia* Fakultas Ilmu Komputer UI.
14. Kepada Bapak/Ibu dosen-dosen penguji nantinya.
15. Semua pihak yang baik secara langsung maupun tidak langsung turut membantu dalam pelaksanaan kerja praktik ini.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan dari semua pihak yang telah membantu. Semoga laporan Tugas Akhir (TA) ini membawa manfaat bagi pengembangan ilmu.

Depok, 8 Juli 2008

Penulis

ABSTRAK

Nama : Femphy Pisceldo

Jurusan : Ilmu Komputer

Judul : Pengembangan Lanjut Pengurai Morfologi untuk Bahasa Indonesia
Berbasiskan Morfologi Dua Tingkat Menggunakan XFST dan LEXC

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.

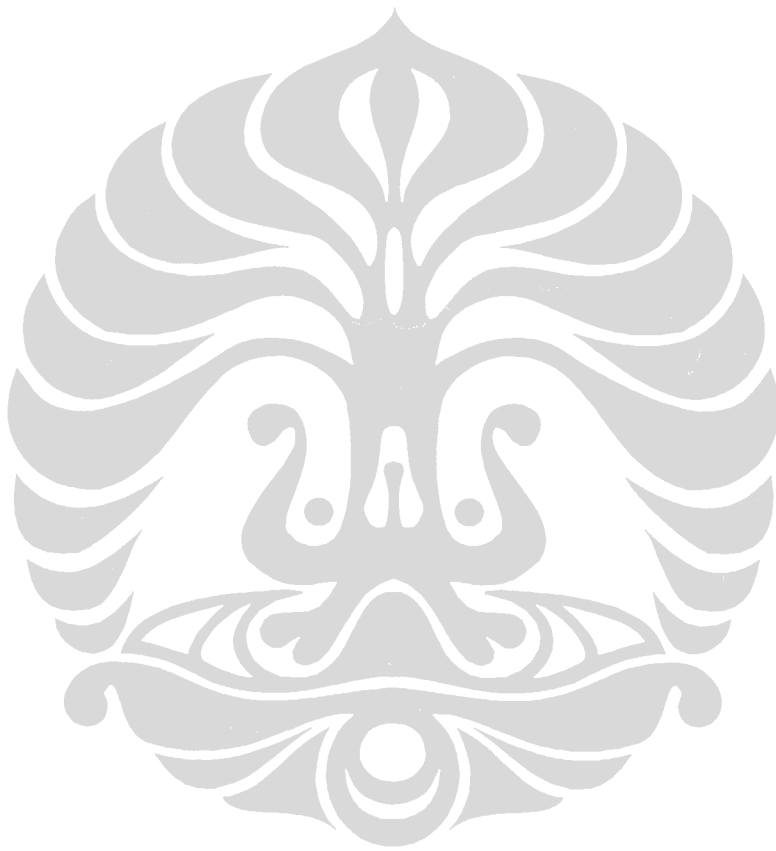
Morfologi bahasa Indonesia yang dicakup dalam penelitian ini berkisar antara kata dasar, kata imbuhan dari penggabungan kata dasar dengan awalan (prefiks), akhiran (sufiks), maupun gabungan pengimbuhan (konfiks), hingga kata ulang sejati, kata ulang sebagian, dan kata ulang berimbuhan. Morfologi bahasa 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.



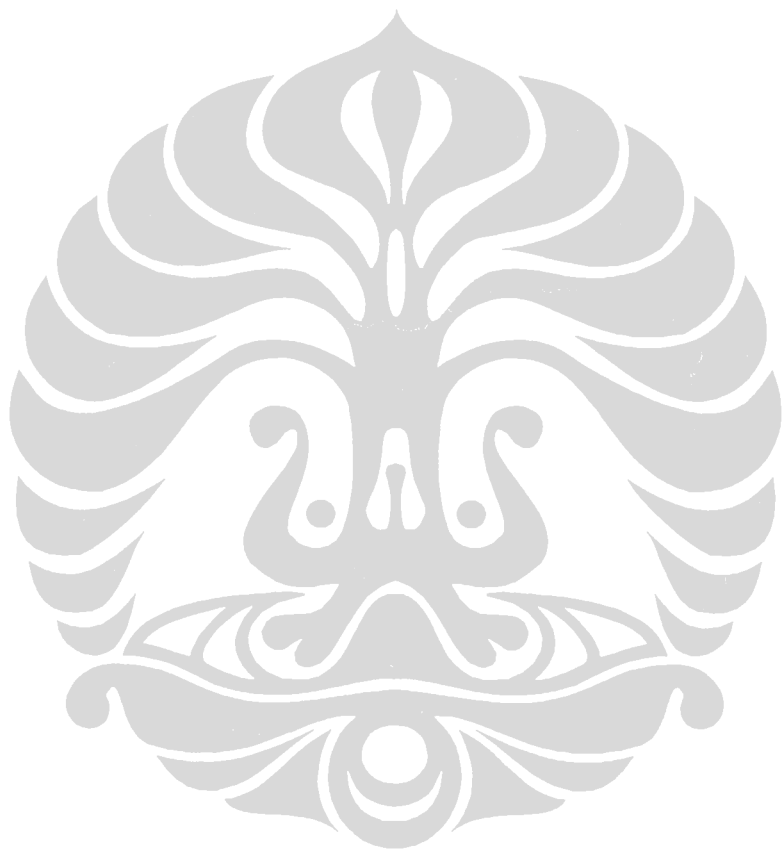
DAFTAR ISI

HALAMAN PENGESAHAN	ii
KATA PENGANTAR	iii
ABSTRAK	v
DAFTAR ISI	vii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xiii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Ruang Lingkup dan Batasan Penelitian	2
1.5 Metode Penelitian	3
1.6 Sistematika Penulisan	3
BAB 2 LANDASAN TEORI	6
2.1 Morfologi	6
2.2 Pengurai Morfologi	7
2.3 <i>Finite-State Transducers</i>	9
2.4 Morfologi Dua Tingkat (<i>Two-Level Morphology</i>)	14
2.5 <i>Nonconcatenative Morphology</i>	15
2.6 Morfologi Bahasa Indonesia	18
2.6.1 Kata Dasar	19
2.6.2 Imbuhan (Afiks)	19
2.6.2.1 Awalan (Prefiks)	19
2.6.2.2 Sisipan (Infiks)	20
2.6.2.3 Akhiran (Sufiks)	20
2.6.2.4 Gabungan Imbuhan (Konfiks)	20

2.6.3 Kata Ulang (Reduplikasi)	20
2.7 Pengurai Morfologi Bahasa Indonesia	21
BAB 3 RANCANGAN PENGURAI MORFOLOGI BAHASA INDONESIA ...	23
3.1 Rancangan <i>Two-Level Morphology</i>	23
3.2 Rancangan <i>Lexicon</i>	24
3.2.1 Kelas Kata Verba.....	25
3.2.2 Kelas Kata Nomina.....	25
3.2.3 Kelas Kata Adjektiva.....	25
3.2.4 Kelas Kata Dll	26
3.3. Rancangan <i>Tags</i>	26
3.4. Aturan-aturan Morfotaktik.....	28
3.4.1 Morfotaktik Imbuhan meN-.....	28
3.4.2 Morfotaktik Imbuhan di-	31
3.4.3 Morfotaktik Imbuhan peN-	34
3.4.4 Morfotaktik Imbuhan per-.....	35
3.4.5 Morfotaktik Imbuhan ke-	37
3.4.6 Morfotaktik Imbuhan ber-.....	38
3.4.7 Morfotaktik Imbuhan ter-	40
3.4.8 Morfotaktik Imbuhan -an.....	40
3.4.9 Morfotaktik Imbuhan -kan.....	41
3.4.10 Morfotaktik Imbuhan -i	42
3.4.11 Morfotaktik Kata Ulang Sejati	42
3.4.12 Morfotaktik Kata Ulang Sebagian.....	44
3.4.13 Morfotaktik Kata Ulang Berimbuhan	45
3.5 Rangkuman Aturan-aturan Morfotaktik	46
3.6 Aturan-aturan Morfofonemik	51
3.6.1 Morfofonemik Prefiks meN-.....	51
3.6.2 Morfofonemik Prefiks peN-.....	54
3.6.3 Morfofonemik Prefiks ber-	56
3.6.4 Morfofonemik Prefiks ter-	57
3.6.5 Morfofonemik Prefiks di-	57

3.6.6 Morfofonemik Prefiks -kan.....	58
3.6.7 Morfofonemik Prefiks -an.....	58
3.6.8 Morfofonemik Prefiks -i.....	59
3.7 Rangkuman Aturan-aturan Morfofonemik Imbuhan	59
BAB 4 IMPLEMENTASI PENGURAI MORFOLOGI BAHASA INDONESIA	63
4.1 <i>Finite-State Authoring Software Tools</i>	63
4.1.1 Xerox Finite-State Tool (XFST)	63
4.1.2 Finite-State Lexicon Compiler (LEXC)	66
4.2 Implementasi Aturan Morfotaktik.....	69
4.2.1 Bagian <i>Root</i>	74
4.2.2 Bagian Prefiks dan Preprefiks	75
4.2.3 Bagian <i>Stems</i>	78
4.2.4 Bagian Sufiks	79
4.2.5 Bagian Pemeriksaan Akhir	82
4.2.6 Morfotaktik Untuk Kasus Kata Ulang.....	83
4.3 Implementasi Aturan Morfofonemik	84
4.3.1 Implementasi Morfofonemik Prefiks meN- dan peN-.....	87
4.3.2 Implementasi Morfofonemik Prefiks ber-, ter-, dan per-	89
4.3.3 Implementasi Morfofonemik Kata Ulang.....	90
4.4 Contoh Kasus untuk Implementasi	90
4.5 Menggabungkan Aturan Morfotaktik dan Morfofonemik	95
BAB 5 HASIL DAN ANALISIS PENGUJIAN	98
5.1 <i>Test Cases</i> Pengujian.....	98
5.2. Hasil Pengujian	100
5.3. Analisis Hasil Pengujian.....	102
BAB 6 PENUTUP	107
6.1 Kesimpulan.....	107
6.2 Hambatan-hambatan yang Ditemui.....	109
6.3 Saran.....	110
6.4 <i>Future Works</i>	110
DAFTAR PUSTAKA	112
LAMPIRAN A	114

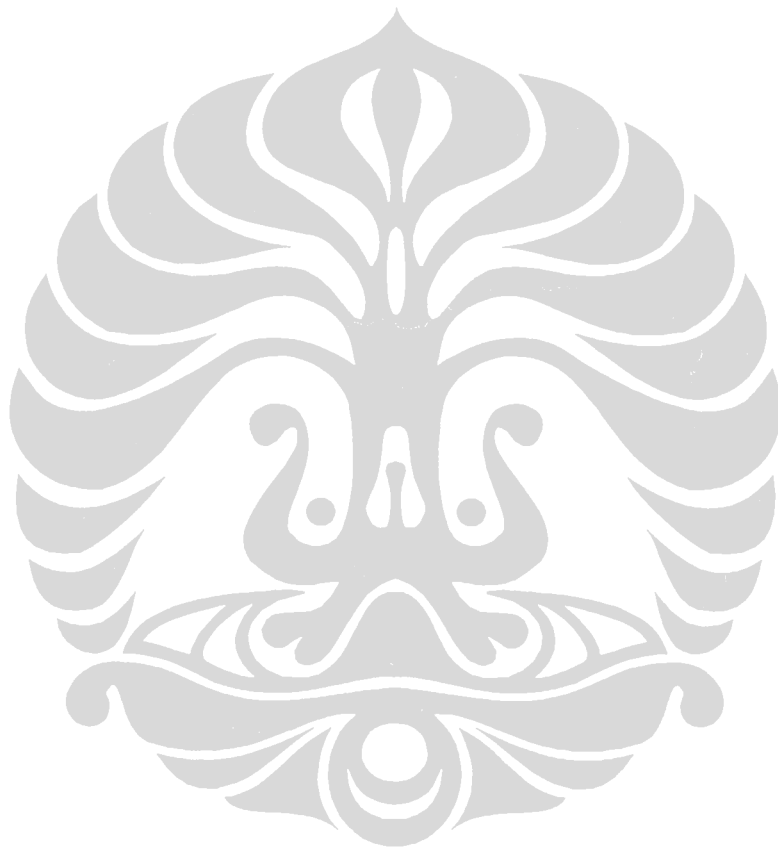
LAMPIRAN B.....	132
LAMPIRAN C.....	186
LAMPIRAN D	197



DAFTAR TABEL

Tabel 3-1. Rancangan <i>Tags</i>	27
Tabel 3-2. Aturan Morfotaktik Imbuhan meN-	29
Tabel 3-3. Aturan Morfotaktik Imbuhan di-	32
Tabel 3-4. Aturan Morfotaktik Imbuhan peN-	34
Tabel 3-5. Aturan Morfotaktik Imbuhan per-	36
Tabel 3-6. Aturan Morfotaktik Imbuhan ke-	37
Tabel 3-7. Aturan Morfotaktik Imbuhan ber-	38
Tabel 3-8. Aturan Morfotaktik Imbuhan ter-	40
Tabel 3-9. Aturan Morfotaktik Imbuhan -an	41
Tabel 3-10. Aturan Morfotaktik Imbuhan -kan	41
Tabel 3-11. Aturan Morfotaktik Imbuhan -i	42
Tabel 3-12. Aturan Morfotaktik Kata Ulang Sejati	43
Tabel 3-13. Aturan Morfotaktik Kata Ulang Sebagian	44
Tabel 3-14. Aturan Morfotaktik Kata Ulang Berimbuhan	45
Tabel 3-15. Variasi yang Diterima dalam Rancangan Morfotaktik	48
Tabel 3-16. Aturan Morfofonemik Prefiks meN-	51
Tabel 3-17. Aturan Morfofonemik Prefiks peN-	54
Tabel 3-18. Aturan Morfofonemik Prefiks ber-	56
Tabel 3-19. Aturan Morfofonemik Prefiks ter-	57
Tabel 3-20. Aturan Morfofonemik Prefiks di-	57
Tabel 3-21. Aturan Morfofonemik Prefiks -kan	58
Tabel 3-22. Aturan Morfofonemik Prefiks -an	58
Tabel 3-23. Aturan Morfofonemik Prefiks -i	59
Tabel 3-24. Tabel Perubahan Fonem Imbuhan	59
Tabel 3-25. Tabel Perubahan Fonem Kata Dasar	61
Tabel 5-1. Tabel Rangkuman Hasil Pengujian untuk Kasus Morfotaktik yang Sah	102
Tabel 5-2. Tabel Rangkuman Hasil Pengujian untuk Kasus Morfotaktik yang Tidak Sah	103
Tabel 5-3. Tabel Rangkuman Hasil Pengujian untuk Kasus Morfofonemik yang Sah	104
Tabel 5-4. Tabel Rangkuman Hasil Pengujian untuk Kasus Morfofonemik yang Tidak Sah	105
Tabel 5-5. Tabel Rangkuman Hasil Pengujian Keseluruhan	106
Tabel A-1. Tabel <i>Test Cases</i> untuk Kasus Morfotaktik yang Sah	114
Tabel A-2. Tabel <i>Test Cases</i> untuk Kasus Morfotaktik yang Tidak Sah	125

Tabel A-3. Tabel <i>Test Cases</i> untuk Kasus Morfofonemik yang Sah	127
Tabel A-4. Tabel <i>Test Cases</i> untuk Kasus Morfofonemik yang Tidak Sah.....	130
Tabel B-1. Tabel Hasil Pengujian untuk Kasus Morfotaktik yang Sah.....	132
Tabel B-2. Tabel Hasil Pengujian untuk Kasus Morfotaktik yang Tidak Sah....	170
Tabel B-3. Tabel Hasil Pengujian untuk Kasus Morfofonemik yang Sah.....	173
Tabel B-4. Tabel Hasil Pengujian untuk Kasus Morfofonemik yang Tidak Sah	184



DAFTAR GAMBAR

Gambar 2-1. Contoh Ilustrasi Skema Pengurai Morfologi	8
Gambar 2-2. Contoh Penggunaan FSA.....	10
Gambar 2-3. Tiga Hasil Keluaran “kammat”	12
Gambar 2-4. Contoh Penggunaan <i>Transducer</i>	13
Gambar 2-5. Ilustrasi Contoh Hubungan <i>Lexical-Surface</i> pada Aturan Dua Tingkat	14
Gambar 2-6. Ilustrasi Contoh Aplikasi Paralel Morfologi Dua Tingkat	15
Gambar 2-7. <i>Chomsky Hierarchy</i>	16
Gambar 2-8. Contoh Skema Pengurai Morfologi untuk <i>Nonconcatenative Morphology</i>	17
Gambar 2-9. Contoh <i>Network Input Compile-Replace</i>	18
Gambar 2-10. Contoh <i>Network Output Compile-Replace</i>	18
Gambar 3-1. Rancangan Pengurai Morfologi	23
Gambar 3-2. Skema Penggabungan <i>Rules</i>	24
Gambar 3-3. Rangkuman Rancangan Morfotaktik.....	47
Gambar 4-1. Cuplikan Code Sederhana dalam Lexc.....	67
Gambar 4-2. Cuplikan Contoh Sederhana Penggunaan <i>Flag Diacritics</i>	68
Gambar 4-4. Cuplikan Pengimplementasian untuk Kasus meNmakan (1).....	70
Gambar 4-3. Contoh <i>Input-Output</i> Aturan Morfotaktik	70
Gambar 4-5. Cuplikan Pengimplementasian untuk Kasus meNmakan (2).....	71
Gambar 4-6. Cuplikan Pengimplementasian untuk Kasus meNmakan (3).....	72
Gambar 4-7. Gambaran Umum Aturan Morfotaktik	73
Gambar 4-8. Cuplikan Implementasi <i>Class Root</i>	74
Gambar 4-9. Cuplikan Implementasi <i>Class</i> Prefiks dan Preprefiks (1).....	75
Gambar 4-10. Cuplikan Implementasi <i>Class</i> Prefiks dan Preprefiks (2).....	76
Gambar 4-11. Ilustrasi Interaksi Prefiks dan Preprefiks	77
Gambar 4-12. Cuplikan Implementasi Bagian <i>Stems</i>	78
Gambar 4-13. Cuplikan Implementasi Bagian Sufiks (1).....	79
Gambar 4-14. Cuplikan Implementasi Bagian Sufiks (2).....	80
Gambar 4-15. Cuplikan Implementasi Bagian Sufiks (3).....	81
Gambar 4-16. Cuplikan Implementasi Bagian Pemeriksaan Akhir	82
Gambar 4-17. Contoh <i>Input-Output</i> Aturan Morfofonemik (1).....	85
Gambar 4-19. Cuplikan Implementasi Contoh pada Gambar 4-18	86
Gambar 4-18. Contoh <i>Input-Output</i> Aturan Morfofonemik (2).....	86
Gambar 4-20. Cuplikan Imlemenatsi untuk Aturan yang Harus Di- <i>compose</i>	87
Gambar 4-21. Cuplikan Implementasi “ <i>^N</i> ” <i>replacement with “n”</i>	87

Gambar 4-22. Cuplikan Implementasi “^N” <i>replacement with “m”</i>	87
Gambar 4-23. Cuplikan Implementasi “^N” <i>deletion</i> khusus peng-	88
Gambar 4-24. Cuplikan Implementasi “e” <i>addition</i> khusus meng-	88
Gambar 4-25. Cuplikan Implementasi “^N” <i>deletion</i> yang diparalel dengan empat aturan perubahan fonem	89
Gambar 4-26. Cuplikan Implementasi “r” <i>deletion</i>	89
Gambar 4-27. Cuplikan Implementasi Penggabungan Semua Aturan Morfofonemik	89
Gambar 4-28. Cuplikan Tambahan untuk Kata Ulang	90
Gambar 4-29. Skema untuk Contoh Kasus Memakan	91
Gambar 4-30. Contoh Kasus Memakan Bagian <i>Root</i>	91
Gambar 4-31. Contoh Kasus Memakan Bagian Preprefiks dan Prefiks	92
Gambar 4-32. Contoh Kasus Memakan Bagian <i>Stems</i>	93
Gambar 4-33. Contoh Kasus Memakan Bagian Redup1	93
Gambar 4-34. Contoh Kasus Memakan Bagian Sufiks	94
Gambar 4-35. Contoh Kasus Memakan Bagian Redup2	94
Gambar 4-36. Contoh Kasus Memakan Bagian Pemeriksaan Akhir	95
Gambar 4-37. Contoh Kasus Memakan Bagian Aturan Morfofonemik	95

