

# PERANGKAT LUNAK PENGUCAPAN KATA BAHASA INDONESIA BERDASARKAN PEMENGGALAN KATA DENGAN FINITE STATE AUTOMATA

**Imam Kuswardayan, Aris Tjahyanto, Krisma Marinda Sormin Siregar**

Jurusan Teknik Informatika,

Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember

Kampus ITS, Jl. Raya ITS, Sukolilo – Surabaya 60111, Telp. + 62 31 5939214, Fax. + 62 31 5913804

Email : imam@its-sby.edu

## ABSTRAK

*Pada tugas akhir ini dikembangkan sebuah perangkat lunak yang berfungsi untuk mengetahui bagaimana pengucapan kata bahasa Indonesia. Pengucapan kata berdasarkan hasil pemenggalan kata dengan menggunakan Finite State Automata (FSA), yaitu suatu model dari mesin pengenali yang mampu mengenali kelas bahasa yang paling sederhana yaitu kelas bahasa Reguler. Pemenggalan kata berdasarkan bagaimana kata tersebut diucapkan. Teks bahasa Indonesia akan disajikan dalam bentuk suara atau ucapan. Pembacaan teks berdasarkan hasil pemenggalan kata. Suara yang disimpan adalah berupa rekaman dari tiap penggalan kata, sehingga akan lebih efisien karena setiap penggalan kata dapat digunakan dalam berbagai kata.*

*Pembuatan perangkat lunak ini didahului dengan perancangan sistem aplikasi. Kemudian rancangan tersebut diimplementasikan. Implementasi aplikasi ini terdiri dari empat proses, yaitu proses pengisian array dari database, proses membersihkan teks, proses FSA dan proses pemenggalan kata.*

*Setelah diimplementasikan, perangkat lunak ini diuji coba sesuai dengan spesifikasi kebutuhan dan kemampuan yang dimiliki yaitu melakukan pemenggalan kata bahasa Indonesia. Seperti memenggal kata dalam sebuah kalimat, kata homograf, diftong, kata berimbuhan dan kata serapan. Hasil pemenggalan kata disajikan dalam bentuk suara atau ucapan. Selain itu, aplikasi ini juga dapat memenggal angka yang sebelumnya dikonversi dulu menjadi kata yang kemudian angka akan dibaca satu per satu. Dengan demikian perangkat lunak ini dapat digunakan sebagaimana mestinya*

**Kata Kunci:** FSA, Pemenggalan Kata.

## 1. PENDAHULUAN

Perkembangan teknologi komputer yang sangat pesat, memicu perkembangan di berbagai bidang yang salah satu diantaranya adalah Teknologi Bahasa Manusia (TBM). TBM pada dasarnya dapat dikelompokkan ke dalam dua kelompok utama yaitu Sistem Bahasa Ucapan (SBU) dan Sistem Bahasa Tertulis (SBT). Dalam kelompok pertama yaitu SBU, salah satu kegiatan yang dilakukan adalah pengembangan sistem pembangkit ucapan dari suatu teks (text-to-speech synthesis system) [4].

Kemajuan pengembangan sistem pembangkit ucapan dari suatu teks jauh tertinggal bila dibandingkan dengan SBT. Beberapa kendalanya adalah pendefinisian sifat-sifat SBU yang merupakan interaksi antara manusia dengan mesin. Hambatan kedua adalah kurangnya alat bantu yang mudah diakses serta murah untuk mendukung riset dan transfer teknologi. Seperti kita ketahui bahwa sistem pembangkit ucapan dari suatu teks merupakan aktivitas yang kompleks.

Mungkin kita berfikir apa gunanya mempelajari tentang pemenggalan kata?

- jika kita melihat peranan pemenggalan kata dalam bahasa tulisan, maka pemenggalan kata perlu dilakukan ketika kata yang kita tulis panjangnya melebihi batas kanan kertas,

- jika kita melihat peranan pemenggalan kata dalam bahasa lisan, maka pemenggalan kata diperlukan untuk mengetahui bagaimana cara mengucapkan suatu kata.

Pada tugas akhir ini, pemenggalan kata ditujukan untuk mengetahui bagaimana pengucapan suatu kata bahasa Indonesia. Sehingga pemenggalan kata berdasarkan bagaimana kata tersebut diucapkan.

Diharapkan tugas akhir ini dapat membantu pengembangan beberapa aspek kebahasaan yang akhirnya dapat membantu pengembangan bahasa Indonesia itu sendiri.

Permasalahan yang dihadapi dalam pembuatan perangkat lunak ini adalah :

1. Bagaimana memenggal kata berdasarkan pengucapannya dalam bahasa Indonesia.
2. Bagaimana hasil pemenggalan kata disajikan dalam bentuk suara.

3. Bagaimana membuat database suara yang mewakili setiap penggalan kata bahasa Indonesia.

Untuk lebih memperjelas dan mencapai tujuan utama pembuatan perangkat lunak ini, maka Tugas Akhir ini dibatasi oleh hal-hal berikut:

- Dalam bahasa Indonesia terdapat kata-kata yang penulisannya sama tetapi pengucapannya berbeda, misalnya kata "apel". Apel dengan huruf "e" yang dibaca seperti kata "perang" bermakna buah apel tapi jika apel dengan huruf "e" yang dibaca seperti pada kata "perak" bermakna upacara atau pertemuan. Hal ini disebabkan karena dalam penulisan menggunakan symbol huruf "e" yang sama untuk suara yang berbeda sehingga dalam tugas akhir ini kata-kata tersebut dianggap sama.
- Menganggap sama untuk kata-kata yang penulisannya sama tetapi memiliki makna kata yang berbeda. Contoh: Beruang.
- Pada kata yang berimbuhan, pemenggalan yang dilakukan tidak berdasarkan pada kata dasarnya. Contoh: Berarti, beriman, berasal.
- Kata-kata seperti 'kacau', 'pantai', 'sepoi' adalah contoh kata yang mengandung diftong. Kata 'kaidah', 'yaitu', 'bau' adalah contoh kata yang tidak mengandung diftong. Karena ketidakkonsistenan ini maka tidak mungkin mengenali diftong secara sempurna tanpa melakukan analisis semantik. Analisis semantik tidak dapat dimodelkan dengan FSA. Sehingga pengenalan diftong dibatasi pada hal-hal yang dapat dimodelkan oleh FSA. Semua pasangan vokal 'au', 'ai' dan 'oi' akan dianggap sebagai diftong.
- Untuk bahasa Indonesia yang diambil dari bahasa asing, ada beberapa kata yang tidak bisa dipenggal berdasarkan pengucapan bahasa Indonesia. Misal, kata transmigrasi.
- Membaca angka satu per satu.

## 2. TEORI OTOMATA

Automata berasal dari bahasa Yunani *automatos*, yang berarti sesuatu yang bekerja secara otomatis (mesin). Dalam hal ini akan dipergunakan istilah *automaton* sebagai bentuk tunggal dan *automata* sebagai bentuk jamak. Teori Automata adalah teori tentang mesin abstrak yang [5]:

- bekerja sekuensial
- menerima input
- mengeluarkan output

Pengertian mesin disini, bukan hanya mesin elektronis/mekanis saja melainkan segala sesuatu (termasuk perangkat lunak) yang memenuhi ketiga ciri di atas. Penggunaan automata pada perangkat

lunak terutama pada pembuatan kompilier bahasa pemrograman.

Secara garis besar ada dua fungsi automata dalam hubungannya dengan bahasa, yaitu:

- Fungsi automata sebagai pengenalan (*RECOGNIZER*) string-string dari suatu bahasa, dalam hal ini bahasa sebagai masukan dari automata.
- Fungsi automata sebagai pembangkit (*GENERATOR*) string-string dari suatu bahasa, dalam hal ini bahasa sebagai keluaran dari automata.

Untuk mengenali string-string dari suatu bahasa, akan dimodelkan sebuah automaton yang memiliki komponen sebagai berikut:

- Pita masukan, yang menyimpan string masukan yang akan dikenali;
- Kepala pita (*tape head*), untuk membaca/menulis ke pita masukan;
- Finite State Controller (FSC), yang berisi status-status dan aturan-aturan yang mengatur langkah yang dilakukan oleh automaton berdasarkan status setiap saat dan simbol masukan yang sedang dibaca oleh kepala pita;
- Peningkat (*memory*), untuk tempat penyimpanan dan pemrosesan sementara.

Automaton pengenalan, setelah membaca string masukan dan melakukan langkah-langkah pemrosesan yang diperlukan, akan mengeluarkan keputusan apakah string tersebut dikenali atau tidak [5].

Pada tugas akhir ini akan menggunakan Finite State Automata sebagai mesin pengenalan untuk bahasa reguler, yaitu kelas bahasa yang paling sederhana. Secara singkat mekanisme FSA dapat dideskripsikan sebagai berikut [5]:

Sebuah string yang terdiri dari satu atau lebih karakter akan menjadi obyek dari sebuah FSA. String tersebut akan dibaca karakter per karakter dimulai dari karakter paling kiri sampai dengan karakter paling kanan. Hasil pembacaan karakter tersebut akan menentukan urutan/rangkaian state yang dikunjungi (dimana awal kunjungan adalah selalu state yang menjadi start state) dan apabila kunjungan tersebut berakhir pada state yang menjadi final state maka string yang dibaca dikatakan "diterima oleh FSA tersebut" atau *accepted*. Sebaliknya jika kunjungan tersebut berakhir pada state selain final state maka dikatakan string tersebut "tidak diterima oleh FSA tersebut" atau *rejected*.

## 3. KONSEP BAHASA INDONESIA

Bahasa Indonesia mengenal bahasa tulisan maupun bahasa lisan. Kadangkala terdapat beberapa perbedaan dalam kedua jenis bahasa ini. Dalam bahasa lisan, dikenal istilah *fonem*, yang merupakan kesatuan bahasa terkecil yang dapat membedakan arti. Dalam bahasa tulisan, *fonem* dilambangkan

dengan huruf. Dengan kata lain, huruf adalah tulisan dari *fonem*. Seringkali istilah fonem disamakan dengan huruf, padahal tidak selamanya berlaku demikian.

Abjad yang digunakan dalam Bahasa Indonesia terdiri atas 52 huruf, yaitu 26 huruf besar (A sampai dengan Z) dan 26 huruf kecil (a sampai dengan z). Huruf vokal dalam ejaan bahasa Indonesia terdiri dari: a, e, i, o dan u. Huruf e melambangkan dua bunyi, yaitu e *taling* dan e *pepet*. Diftong dieja dengan gabungan dua huruf vokal yang dilafazkan sebagai bunyi luncuran. Dalam bahasa Indonesia terdapat tiga jenis huruf diftong yang dieja dengan gabungan huruf vokal sebagai berikut: ai, au, dan oi. Tetapi ada beberapa kata yang mengandung gabungan dua huruf vokal yang bukan merupakan diftong, seperti kaidah, bau dan yaitu. Huruf konsonan dalam ejaan bahasa Indonesia terdiri dari huruf-huruf b, c, d, f, g, h, j, k, l, m, n, p, q, r, s, t, v, w, x, y, dan z. Dalam bahasa Indonesia terdapat empat gabungan huruf konsonan, yaitu kh, ng, ny, dan sy, yang masing-masing melambangkan satu bunyi konsonan. Bahasa Indonesia mengenal beberapa pola umum penggalan kata, yaitu V tunggal, VK, KV, KVK, KKV, KKVK, VKK, KVKK, KKVKK, KKKV, KKKVK.

Untuk memenggal kata, dapat digunakan pedoman berikut ini:

- Kalau di tengah kata terdapat dua vokal berturutan (selain diftong), pemisahan dilakukan di antara kedua vokal tersebut.  
Contoh: siapa → si – a – pa
- Kalau di tengah kata terdapat konsonan di antara dua vokal, pemisahan dilakukan sebelum konsonan tersebut.  
Contoh : aku → a – ku
- Kalau di tengah kata terdapat dua konsonan atau lebih, pemisahan dilakukan setelah konsonan pertama.  
Contoh : rinda → rin – da
- Kalau di tengah kata ada tiga konsonan atau lebih, pemisahan tersebut dilakukan diantara huruf konsonan yang pertama (termasuk gabungan huruf konsonan) dengan huruf konsonan yang kedua.  
Contoh: kontra → kon-tra, angklung → ang-klung
- Imbuhan dan partikel yang biasanya ditulis serangkai dengan kata dasar, pada penyukuan dipisahkan.  
Contoh : persekutuan → per – se – ku – tu – an

Pedoman a sampai dengan d, bisa diterapkan pada perancangan perangkat lunak ini. Pedoman e tidak bisa diterapkan, karena pedoman tersebut lebih tepat diterapkan pada bahasa tulisan. Pada bahasa lisan tidak dikenal imbuhan maupun partikel, yang ada hanyalah bagaimana kata tersebut diucapkan. Pemenggalan kata ditujukan untuk mengetahui bagaimana pengucapan suatu kata bahasa Indonesia.

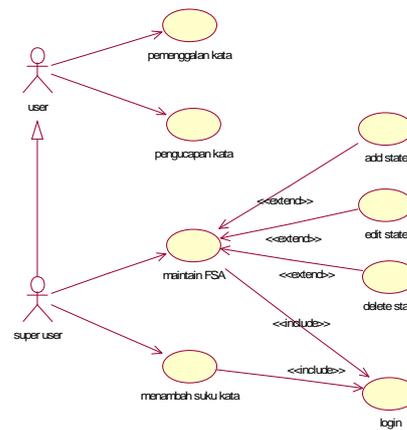
Dalam hal ini, penggalan kata dianggap sebagai satuan bahasa lisan. Hal ini perlu dipertegas karena kadangkala pemenggalan kata dalam bahasa lisan dan bahasa tulisan berbeda [6].

#### 4. PERANCANGAN APLIKASI

Perancangan aplikasi yang perlu dilakukan meliputi perancangan interface, perancangan proses, perancangan Finite State Automata dan dan perancangan Database. Perancangan interface berguna untuk memudahkan dalam membuat aplikasi nantinya. Perancangan interface terdiri dari perancangan form utama, form pengeditan aturan Finite State Automata dan form Login.

Perancangan proses berguna untuk mengintegrasikan semua proses yang terjadi dalam aplikasi yang akan dibuat. Perancangan proses merupakan gambaran fisik bagaimana aplikasi ini bekerja. Inti dari proses dalam sistem aplikasi ini adalah dimulai dengan pengguna memasukkan teks bahasa Indonesia. Kemudian dengan menggunakan Finite State Automata teks yang diinputkan akan dikenali dan dipenggal sesuai pengucapan kata. Penggalan kata akan disuarakan dengan ucapan bahasa Indonesia.

Pembuatan rancangan proses ini menggunakan software Rational Rose 2000. Proses yang terjadi digambarkan dengan beberapa diagram, yaitu *Use Case diagram*, *Sequence Diagram* dan *Class Diagram*. Berikut adalah use case diagram yang menggambarkan proses-proses pada sistem aplikasi:



Gambar 1. Use Case Diagram

Perancangan Finite State Automata adalah perancangan aturan pemenggalan kata. Dalam perancangan perangkat lunak ini digunakan Finite State Automata sebagai mesin abstrak yang akan mengenali dan memenggal kata dalam kalimat. FSA yang akan digunakan dirancang dalam tiga tingkatan. Pada tingkatan pertama yang akan dikenali adalah pola-pola :V,K atau KV. Hasil pengenalan FSA pada suatu tingkatan menjadi masukan bagi FSA tingkatan berikutnya. Pada tingkatan kedua FSA akan mengenali penggalan kata dengan pola V, VK, VKK,

KV, KVK, KKV, KKVK, KKKV, KKKVK. Begitu pula pada FSA tingkat ketiga akan menerima input dari output FSA tingkat kedua.

Dari kelakuan FSA tingkatan kedua, tampak bahwa pola penggalan kata VKK, KVKK dan KKVKK belum bisa dikenali. Untuk itu diperlukan FSA tingkatan ketiga yang mampu mengenalinya. Jika FSA ini menemukan pola VK-K, KVK-K dan KVKK-K, maka akan dikenali sebagai pola penggalan kata VKK, KVKK dan KKVKK. Selain itu FSA pada tingkatan 3 dapat mengenali diftong.

Perancangan database adalah perancangan atau desain tabel-tabel yang berguna untuk menyimpan data-data dan aturan-aturan FSA yang digunakan dalam sistem aplikasi.

Dalam database ini terdapat empat tabel, yaitu tabel *Finite State Automata (FSA)*, tabel *Pengkodean Pola*, tabel *Login* dan tabel *Suku kata*.

Tabel FSA berguna untuk menyimpan aturan-aturan (rule) dari proses Finite State Automata. Tabel *Pengkodean Pola* berguna untuk menyimpan jenis karakter atau pola-pola dari pemenggalan kata menjadi. Tabel *Login* adalah tabel yang berguna untuk menyimpan data pengguna yang memiliki hak khusus atau yang disebut sebagai super user.

Tabel yang keempat adalah tabel *Sukukata*, tabel ini hanya memiliki satu field yaitu field *kata* dengan tipe data teks yang berfungsi untuk menyimpan penggalan kata hasil proses pemenggalan kata oleh sistem aplikasi.

## 5. IMPLEMENTASI PERANGKAT LUNAK

Perangkat lunak ini diimplementasikan dengan menggunakan Microsoft Access 2000 untuk membuat database serta Visual Basic 6.0 sebagai bahasa pemrograman.

Implementasi aplikasi ini meliputi tiga hal, yaitu implementasi interface, implementasi database dan implementasi proses. Implementasi interface adalah membuat form-form dan menu-menu yang telah dirancang sebelumnya. Implementasi database adalah mengisi data-data yang dibutuhkan atau digunakan untuk proses pemenggalan kata.

Berikut ini implementasi beberapa proses yang ada dalam sistem:

### a) Proses Pengisian Array dari Database

Proses ini terjadi pada saat loading form pertama kali sistem aplikasi dijalankan. Implementasi proses ini adalah sebuah prosedur yang dipanggil hanya sekali pada saat form load. Prosedur ini akan melakukan pengisian array dari database, yaitu untuk proses FSA pada sistem aplikasi.

### b) Proses Membersihkan Teks

Proses membersihkan teks adalah proses membuang karakter-karakter selain huruf/abjad, angka, tanda-tanda operasi (-, +, /, \*, =), dan lambang "&", "%", "\$". Misal, spasi, tanda baca dan lainnya yang tidak perlu dibaca atau disuarakan dalam sebuah teks. Pembersihan teks ini dilakukan sebelum proses

pemenggalan kata dilakukan yaitu sebelum proses FSA tingkat pertama. Fungsi ini akan diproses setiap kali akan dilakukan pemenggalan kata (pada saat user meng-klik tombol *pemenggalan kata*). Karakter-karakter yang tidak dikenali akan diabaikan atau tidak dianggap sebagai input karakter.

### c) Proses FSA

Proses ini merupakan proses pengenalan karakter yang diinputkan. Proses ini akan menyimpan aturan-aturan FSA dari database untuk disimpan kedalam array. Proses ini diimplementasikan dalam sebuah fungsi yang terdiri dari tiga fungsi, yaitu Fungsi *prosesFA1*, fungsi *prosesFA2* dan fungsi *prosesFA3*. Setiap fungsi berisi aturan-aturan dari setiap tingkat FSA.

Fungsi *prosesFA1* akan mengenali pola-pola pada Finite State Automata tingkat pertama. Fungsi *prosesFA1* akan menerima input karakter (berupa string) yang diinputkan setelah dibersihkan dan fungsi ini secara keseluruhan akan menentukan *next state* dari proses FSA, yaitu dengan karakter yang diinputkan maka akan menuju state yang sesuai, dimana *next state* merupakan array yang tersimpan dalam database dan digunakan pada proses selanjutnya.

Fungsi *prosesFA2* merupakan proses FSA tingkat kedua, yaitu mengenali pola-pola pada Finite State Automata tingkat kedua. Fungsi ini pada dasarnya sama dengan fungsi *prosesFA1*, hanya saja input yang diterima merupakan hasil penguraian teks pada FSA tingkat pertama. Fungsi ini juga menentukan *next state* dari proses FSA sesuai dengan pola karakter atau pola penggalan kata yang terdapat pada FSA tingkat kedua.

Fungsi *prosesFA3* akan mengenali pola-pola pada Finite State Automata tingkat ketiga. Fungsi ini juga pada sama dengan fungsi *prosesFA2*, menerima input dari proses sebelumnya, dalam hal ini fungsi *prosesFA3* akan menerima input dari hasil penguraian teks pada FSA tingkat kedua.

Proses ini merupakan proses FSA yang terakhir, dimana proses ini akan dipanggil pada proses pemenggalan kata pada FSA tingkat 3 yang menghasilkan penggalan kata hasil pemenggalan teks secara keseluruhan.

### d) Proses pemenggalan kata

Proses ini bertujuan untuk melakukan pemenggalan kata. Proses ini diimplementasikan dalam fungsi (function). Terdapat tiga fungsi sesuai dengan proses FSA yang memiliki tiga tingkat, yaitu fungsi *PecahdgnFA1* untuk FSA tingkat pertama, fungsi *PecahdgnFA2* untuk FSA tingkat kedua, fungsi *PecahdgnFA3* untuk FSA tingkat ketiga.

Fungsi *PecahdgnFA1* menerima input berupa teks yang dimasukkan oleh user, sebelum dilakukan pemenggalan kata, teks akan dibersihkan seperti yang sudah dijelaskan dengan begitu Fungsi *PecahdgnFA1* akan memanggil *Function siapkantextuntukdiproses*. Setelah teks dibersihkan fungsi ini akan memanggil

fungsi *prosesFA1* berkali-kali selama *NextState = prosesFA1*. Kemudian kata-kata dalam teks akan dipenggal dimana penggalan kata merupakan hasil pecahan dari Finite State Automata tingkat pertama.

Seperti halnya FSA tingkat pertama, FSA tingkat kedua yaitu fungsi *PecahdgnFA2* akan memecah/menguraikan kembali hasil dari pemenggalan FSA tingkat pertama. Pemenggalan sesuai dengan pola-pola yang terdapat pada FSA tingkat kedua. Proses ini akan memanggil fungsi *prosesFA2* dan memprosesnya selama *NextState = prosesFA2*.

Kemudian proses yang terakhir adalah proses pemenggalan kata dengan FSA tingkat 3 yang diimplementasikan dengan fungsi *PecahdgnFA3*. Fungsi ini akan memproses hasil dari pemenggalan FSA tingkat kedua. Fungsi *prosesFA3* akan dipanggil selama *NextState = prosesFA* dan menghasilkan output berupa penggalan kata hasil pecahan sesuai FSA tingkat ketiga.

Proses pemenggalan kata akan berhenti pada FSA tingkat ketiga, karena penggalan kata yang dihasilkan sudah sesuai dengan pengucapan bahasa Indonesia. Penyuaan teks atau pembacaan teks berdasarkan penggalan kata hasil dari proses pemenggalan kata FSA tingkat ketiga.

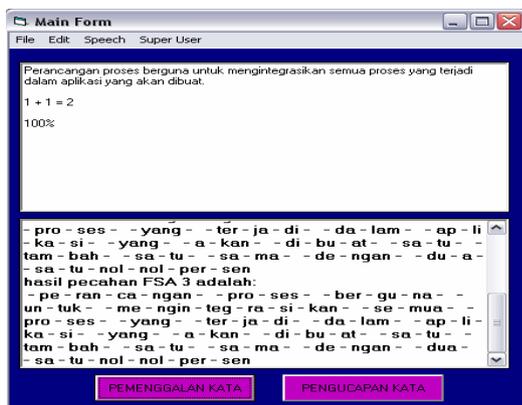
## 6. HASIL UJI COBA

Pada uji coba ini diberikan dua buah skenario yang ditujukan untuk mengetahui fungsionalitas dari aplikasi yang dibuat, yaitu :

### 1. Pemenggalan kata dan angka

Pada ujicoba kali ini akan dicoba untuk memasukkan sebuah kalimat dan sebuah operasi matematika yaitu penambahan serta nilai "100%".

Aplikasi akan memenggal setiap kata dalam kalimat. Kemudian angka dan tanda oprasi juga akan diuraikan begitu pula lambang persen. Angka dan lambang-lambang atau tanda operasi akan dikonversi terlebih dahulu menjadi teks kemudian baru dilakukan pemenggalan kata. Seperti terlihat pada **Gambar 2**.

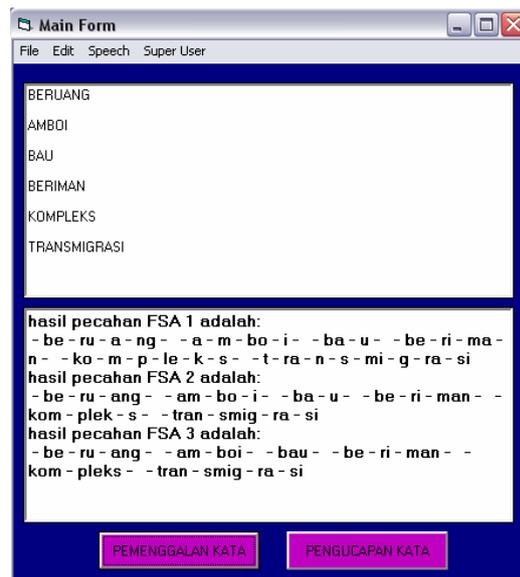


Gambar 2 . Hasil Uji Coba 1

Angka akan dibaca satu per satu , seperti terlihat pada gambar di atas "100" akan dibaca "satu-nol-nol". Hal ini dikarenakan karakter akan dikenali, diuraikan dan kemudian dibaca satu per satu.

### 2. Pemenggalan kata homograf, diftong, gabungan huruf vokal tapi bukan diftong, kata berimbuhan dan kata serapan.

Uji coba yang kedua akan memenggal kata "beruang" (homograf), "amboi" (diftong), "bau" (bukan diftong), "beriman" (kata berimbuhan), "kompleks" dan "transmigrasi" (kata serapan).



Gambar 3. Hasil Uji Coba 2

Kata beruang yang berarti binatang dan kata beruang yang berarti memiliki uang banyak akan diuraikan dan dibaca sama yaitu "be-ru-ang" walaupun memiliki makna yang berbeda. Kata yang merupakan diftong atau yang bukan diftong, penguraiannya sama seperti kata yang merupakan diftong. Untuk membedakannya diperlukan analisis semantik sedangkan analisis semnatik tidak bisa dimodelkan dengan FSA. Kata kompleks adalah kata serapan dapat dipenggal dengan sempurna sesuai dengan bahasa lisannya yaitu "kom-pleks" tetapi kata transmigrasi tidak bisa dipenggal sesuai dengan ucapannya sehingga kata-kata serapan tertentu untuk dapat dipenggal sesuai dengan pengucapannya diperlukan aturan-aturan tambahan.

Secara umum sistem aplikasi ini telah berhasil diimplementasikan dengan baik, sesuai dengan tujuan pembuatan aplikasi yaitu dapat melakukan pemenggalan kata bahasa Indonesia menyajikannya dalam bentuk suara atau ucapan. Dan untuk penyempurnaan aplikasi, proses Finite State Automata dapat dirubah/maintain (desain FSA) langsung pada sistem aplikasi, sehingga tidak perlu membuka database. Selain itu juga untuk memperkaya penggalan kata dalam database,

penambahan dapat dilakukan langsung pada main form sistem aplikasi.

Namun begitu, dari hasil beberapa kali uji coba dan evaluasi terhadap sistem aplikasi ini, masih ditemukan beberapa kekurangan terutama dalam hal pemenggalan kata. Ada beberapa kata bahasa Indonesia yang diambil dari bahasa Asing (kata serapan) yang tidak dapat dipenggal secara sempurna oleh aplikasi sesuai dengan bahasa lisannya atau pembacaan kata seharusnya. Selain itu, aplikasi belum bisa membaca angka yang berupa jumlah nilai uang (ribuan, jutaan, dan sebagainya) jadi inputan berupa angka akan dibaca satu per satu.

Sistem aplikasi ini tidak memfokuskan pada kualitas suara atau intonasi suara, sehingga hasil suara dari pembacaan teks tidak layaknya seperti pembacaan teks secara langsung. Intonasi dari pembacaan teks semuanya dengan intonasi datar. Hal ini disebabkan karena suara direkam dengan intonasi datar atau pada semua penggalan kata memiliki intonasi yang sama.

## 7. KESIMPULAN & SARAN

Beberapa kesimpulan yang dapat diambil dari Tugas Akhir ini adalah :

1. Dengan Finite State Automata dapat dilakukan pemenggalan kata berdasarkan pengucapannya.
2. Dapat mengetahui bagaimana pengucapan kata bahasa Indonesia berdasarkan hasil pemenggalan kata.
3. Dengan menyimpan suara per penggalan kata dalam database, dapat menghemat penyimpanan di dalam sistem  $\pm 35\%$ . Hal ini lebih efisien karena dari penggalan kata dapat digunakan diberbagai kata.

Berikut ini adalah saran-saran yang dapat diberikan untuk pengembangan Tugas Akhir lebih lanjut:

1. Aplikasi ini diharapkan nantinya bisa menjadi kelas library (dd, axtiveX, dab sebagainya).
2. Kata-kata bahasa Indonesia yang diambil dari bahasa Asing semuanya dapat dipenggal secara sempurna sesuai dengan pengucapan (bahasa lisan) dari kata tersebut.
3. Penyuaran teks dapat lebih baik, yaitu kualitas rekaman suara dan intonasi suara lebih sempurna sehingga dapat berguna untuk produk-produk yang memanfaatkan teknologi Text To Speech yang dapat bermanfaat untuk masyarakat.
4. Database suara yang digunakan untuk menyajikan dalam bentuk suara tidak lagi menggunakan file wav, tetapi menggunakan mekanisme perubahan rekuensi suara.

## 7. DAFTAR PUSTAKA

1. Arman, A., Arry. "Perkembangan Teknologi TTS Dari Masa ke Masa". 2003.
2. Basuki, A., Thomas. "Pengenalan Suku Kata Bahasa Indonesia Menggunakan Finite-State Automata". Integral, 2000.
3. Cohen, I. A., Daniel. "Introduction To Computer Theory". John Wiley & Sons, New York Chichester Brisbane Toronto, 1986.
4. Gunarso. "Pemanfaatan Teknologi Dalam Pengembangan Bahasa Indonesia". Direktorat Teknologi Informatika dan Elektronika, Badan Pengkajian dan Penerapan Teknologi, 1998.
5. Hopcroft, J.E.; Ullman, J.D., "Introduction to Automata Theory, Languages and Computation". Addison-Wesley, 1979.
6. Pusat Bahasa Departemen Pendidikan Nasional, "Pedoman Umum Ejaan Bahasa Indonesia Yang Disempurnakan". Balai Pustaka, Jakarta, 2003.
7. Sudarno. "Pemisahan Kata Atas Suku Katanya". Rampak Serantau, 1994.