

Analisis Leksikal untuk Mengidentifikasi Kata Kerja

(Lexical Analysis to Identify a Verb)

Ridho Muktiadi¹⁾, Septian Ari Wibowo²⁾

¹⁾²⁾ *Teknik Informatika – Fakultas Teknik – Universitas Muhammadiyah Purwokerto
Jalan raya Dukuwaluh Po.Box 202 Purwokerto – Jawa Tengah 53182*

¹⁾ *ridho.muktiadi@gmail.com*

²⁾ *septiannim12@gmail.com*

Abstrak— Kalimat terdiri dari berbagai macam kata, dalam pembentukan kalimat terdapat berbagai macam sifat dan kedudukan kata yang digunakan. Salah satunya yaitu kata kerja. Kata kerja dapat terdiri dari kata kerja asli, atau perubahan dari kata sifat dan kata yang lainnya yang memperoleh imbuhan sehingga berubah menjadi sebuah kata kerja. Dalam metode analisis leksikal, kata disebut dengan token, dimana analisis leksikal melakukan scanning terhadap karakter yang sesuai dengan token yang sudah ditetapkan. Token yang diidentifikasi dalam penelitian ini adalah kata kerja, dimana kata kerja terdiri dari kata kerja murni dan kata kerja yang memperoleh imbuhan. Menggunakan tool Visual Studio 2010 dan bahasa pemrograman C# dalam penelitian ini dapat menciptakan sebuah aplikasi yang dapat mendeteksi kata kerja pada kalimat. Menggunakan aplikasi ini kata kerja dapat teridentifikasi sesuai dengan data yang terdapat dalam database.

Keywords— Analisis Leksikal, Identifikasi, kata kerja

Abstract— Sentence consists of various kinds of word, sentence formation contained a wide variety of properties and position of the words used. One of them is a verb. The verb can be composed of the original verb, or a change of adjectives and others which obtain affixes that turned into a verb. In the lexical analysis method, word called by tokens, where lexical analysis scanning the character that match to a predefined tokens. Tokens that identified in this study is a verb, consists of a pure verb and verb that obtain affixes. Using the tools, Visual Studio 2010 and C # programming language in this research can create an application that can detect verb in the sentence. Using this application verb can be identified according to the data contained in the database.

Keywords— Lexical analysis, identification, Verb.

I. PENDAHULUAN

Kalimat adalah kata atau sekelompok kata yang merupakan satu kesatuan yang mengutarakan suatu

pikiran dan perasaan. Kata merupakan apa-apa yang dilahirkan dari ucapan, sedangkan kata kata kerja adalah kata yang menyatakan perbuatan [1]. Penentuan kata kerja pada kalimat dengan menggunakan kinerja system pada perangkat lunak, merupakan salah satu upaya pengembangan dan penerapan ilmu teknologi informasi.

Analisis Leksikal merupakan sebuah metode yang akan diterapkan pada sistem ini, yang merupakan salah satu bidang ilmu Pemrosesan Bahasa Alami yang merupakan bagian dari Sistem Cerdas. Analisis Leksikal dapat mengidentifikasi semua besaran pembangunan bahasa pada kode sumber. Oleh karena itu, dalam penelitian ini diciptakan sebuah aplikasi yang dapat digunakan untuk mendeteksi kata kerja pada kalimat dengan menggunakan metode analisis leksikal. Diharapkan dengan menggunakan metode ini nantinya dapat digunakan untuk mengidentifikasi kata kerja pada kalimat, dan menjadi sebuah pengembangan dari ilmu teknologi informasi khususnya dalam bidang pemrosesan bahasa alami.

II. METODE

A. Jenis Penelitian

Penelitian ini merupakan penelitian pengembangan untuk mengidentifikasi kata kerja pada kalimat dengan menggunakan metode analisis leksikal.

B. Tempat dan Alat

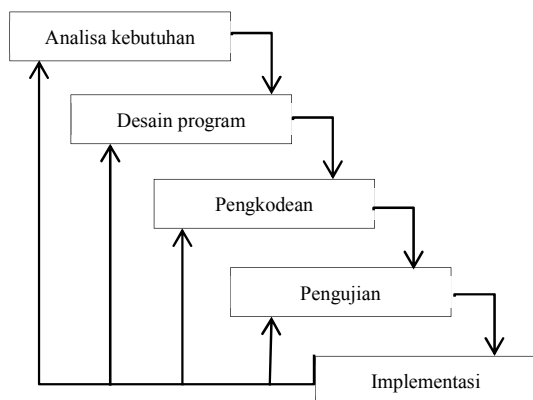
Penelitian ini dilakukan di Universitas Muhammadiyah Purwokerto. Alat untuk mengembangkan sistem ini adalah menggunakan software Visual Studio 2010 dan Ms Sql Server 2005 Express edition. Sedangkan untuk hardwarenya menggunakan laptop dengan spesifikasi processor Intel Core, RAM 2Gb, Hardisk minimal 40Gb.

C. Pengumpulan data

Pengumpulan data dilakukan melalui studi literatur untuk mengetahui tentang metode Analisis Leksikal dan struktur dalam sebuah kalimat.

D. Tahap Pengembangan Sistem

Tahap ini dilakukan meliputi analisa kebutuhan, desain program, pengkodean (*coding*), pengujian, dan implementasi sistem seperti terlihat pada Gambar 1.



Gambar 1. Diagram Alur Pengembangan

1) *Analisis Kebutuhan*: Pada tahap ini dikumpulkan data kata kerja bahasa Indonesia yang diambil dari Kamus Umum Bahasa Indonesia.

1) *Desain*: Program dirancang menggunakan langkah-langkah sebagai berikut [2]:

- menentukan / mendeskripsikan token
- menggunakan *regular grammar*.
- menggunakan ekspresi *grammar*.
- menspesifikasikan token-token dengan ekspresi reguler.
- membuat model matematis menggunakan *finite-state acceptor* (FSA) atau *finite automata* (FA).
- mengimplementasi Analisis Leksikal sebagai Finite Automata (FA)

2) *Pengkodean*: Pengkodean digunakan untuk membangun aplikasi, dalam pembuatan sistem ini adalah dengan bahasa pemrograman C# [3], Visual Studio 2010 dan Structur Query Language (SQL) [4].

3) *Pengujian*: Metode pengujian yang digunakan adalah *black box testing*. *Black box testing* adalah pengujian program yang dilakukan oleh pengembang (*programmer*) dengan memberikan *input* tertentu dan melihat hasil yang didapatkan dari *input* tersebut. Dengan kata lain *black box testing* terfokus pada fungsionalitas sistem [5].

4) *Implementasi*: Pada tahap ini merupakan tahap untuk merealisasikan sistem yang baru dibangun

sehingga dapat menentukan susunan kata kerja pada sekumpulan kalimat.

III. HASIL DAN PEMBAHASAN

A. Penentuan Kebutuhan Data

Data yang digunakan di dalam penelitian ini adalah kata kerja dalam Bahasa Indonesia, yang berasal dari kata tertentu yang memperoleh imbuhan dan akhiran. Namun terdapat juga kata kerja yang tidak memperoleh imbuhan dan akhiran kata.

B. Penyelesaian masalah menggunakan Analisis Leksikal

1) *Menentukan / mendeskripsikan token*: Token yang digunakan berasal dari kata kerja Bahasa Indonesia yang terdiri dari berbagai jenis. Token sendiri terdiri dari beberapa karakter yang masuk dan terpisahkan menjadi *identifiser*, *keyword* dan konstanta. Berikut daftar kata kerja yang tersaji dalam Gambar 2.

Kode Kata	Kata
K0006	bekerja
K0003	berangkat
K0012	berdendang
K0004	berenang
K0010	bertali
K0005	beman
K0009	bemanyari
K0011	bemusa
K0008	beranak

Gambar 2. Daftar kata kerja

2) *Menggunakan regular grammar*: Regular Grammar yang digunakan dalam spesifikasi Analisis Leksikal untuk mendeteksi kata kerja adalah sebagai berikut:

- I → aA|bA|.....|zA|a|b|.....|z,
- Ber → aA|bA|.....|zA,
- Di → aA|bA|.....|zA,
- Ke → aA|bA|.....|zA,
- Meng → aA|bA|.....|zA,
- Per → aA|bA|.....|zA,
- Ter → aA|bA|.....|zA, Ber → aA|bA|.....|zA → Kan,
- Ber → aA|bA|.....|zA → An,
- Di → aA|bA|.....|zA → Kan, Di → aA|bA|.....|zA → i,
- Ke → aA|bA|.....|zA → i, Ke → aA|bA|.....|zA → An,
- Meng → aA|bA|.....|zA → Kan,
- Meng → aA|bA|.....|zA → i,

Per → aA|bA|.....|zA → Kan,
 Per → aA|bA|.....|zA → i,
 Ter → aA|bA|.....|zA → Kan,
 Ter → aA|bA|.....|zA → i,

Reguler gramer di atas disusun berdasarkan pola pembentukan kata kerja yang mempunyai awalan, akhiran, dan juga awalan dan akhiran. Berikut pola pembentukan dari kata kerja (Tabel I).

TABEL I
 TABEL POLA PEMBENTUKAN KATA KERJA

Prefiks/sufiks	-kan	-i	-an
ber-	v		v
di-	v	v	
ke-		v	v
meng-	v	v	
per-	v		
ter-	v	v	

3) *Menggunakan ekspresi grammar*: Ekspresi ini digunakan untuk menspesifikasikan token-token dengan ekspresi reguler. Berikut digunakan ekspresi reguler yang ekuivalen dengan reguler grammar di atas:

$$I = (a|b|...z)(a|b|...z)^* = \text{huruf}(\text{huruf})^*$$

4) *Menspesifikasikan token-token dengan ekspresi reguler*: Dalam spesifikasi sintak biasanya digunakan *context free grammar* (CFG). Sebagai contoh ekspresi if-then, dengan simbol awal E adalah seperti berikut:

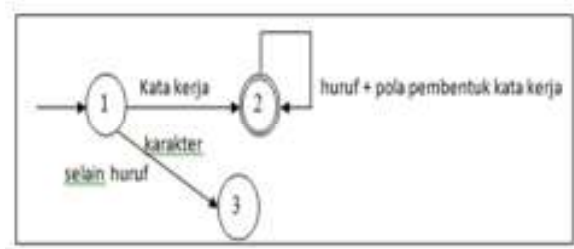
$$E \rightarrow \text{if } L \text{ then, } L \rightarrow IO,$$

$$I = \text{huruf}(\text{huruf})^*,$$

$$O \rightarrow \langle | = | \rangle \langle | = | \rangle = | \dots | 0 | I | \dots | 9.$$

5) *Membuat model matematis menggunakan finite-state acceptor (FSA) atau finite automata (FA)*: Penentuan model matematis menggunakan *finite-state acceptor* (FSA) berdasarkan jenisnya, sehingga digunakan model DFA (*Deterministic Finite Automata*). Model ini membentuk skema *scanner* seperti pada Gambar 3.

Setelah tersusun skema *scanner*, selanjutnya dibuat sebuah aturan translasi untuk beberapa ekspresi reguler atau token.

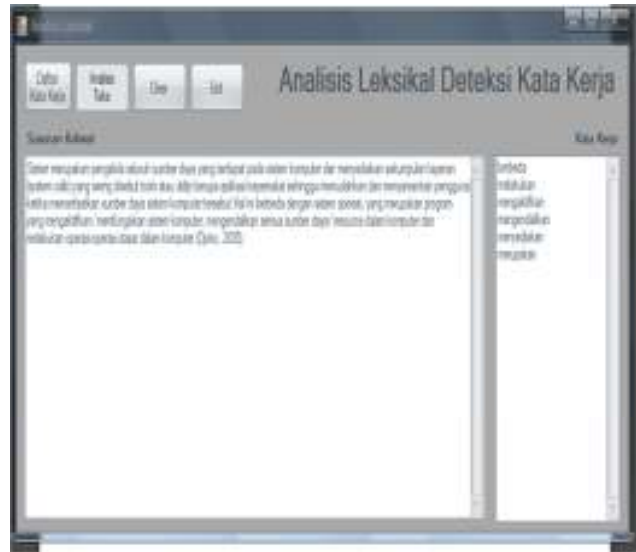


Gambar 3. Skema Scanner Kata Kerja

- jika *state* i merupakan *state* akhir berarti pada *state* i telah ditemukan sebuah token
- jika *state* i bukan *state* akhir berarti pada *state* i telah terdeteksi token tidak dikenal.

6) *Mengimplementasi Analisis Leksikal sebagai Finite Automata (FA)*: Implementasi dari Analisis Leksikal untuk mendeteksi kata kerja dibangun menggunakan Visual Studio 2010 dengan bahasa pemrograman C#. Berikut hasil dari implementasi Analisis Leksikal untuk mendeteksi kata kerja (Gambar 4). Pada gambar tersebut terdapat suatu tempat untuk memasukkan beberapa kalimat, dan sebelah kanan merupakan hasil dari analisis kata kerja pada kalimat yang telah diidentifikasi.

Penelitian ini hanya digunakan untuk mengidentifikasi kata kerja dengan memberikan masukan terlebih dahulu pada database penyimpanan kata kerja tersebut. Sehingga dalam pendeteksian pola kata kerja masih mengacu ke database. Dengan diberikan pola pada sebuah kata kerja, nantinya dapat terdeteksi sebuah kata kerja dimana terdapat kekurangan penulisan karakter pada imbuhan dari kata kerja tersebut.



Gambar 4. Aplikasi Analisis Leksikal untuk mengidentifikasi kata kerja

IV. PENUTUP

A. *Simpulan*

Berdasarkan hasil dan pembahasan dari penelitian, dapat disimpulkan bahwa dengan menggunakan Analisis Leksikal dapat digunakan untuk mengidentifikasi kata kerja pada sebuah kalimat. Dengan masukan berupa kalimat, maka akan teridentifikasi kata kerjanya.

B. *Saran*

Daftar kata-kata yang tersimpan dalam database masih sangat terbatas, sehingga disarankan untuk bisa dilengkapi. Selain itu juga disarankan untuk dikembangkan analisis leksikal untuk mengidentifikasi

jenis kata yang lain misalnya kata sifat, kata benda, dan kata keterangan. Bahkan bisa dikembangkan pula untuk mengidentifikasi struktur kalimat yaitu subyek, predikat, maupun obyek.

DAFTAR PUSTAKA

- [1] Pusat Bahasa Departemen Pendidikan Nasional. 2007. Kamus Umum Bahasa Indonesia. Balai Pustaka. Jakarta.
- [2] <http://www.globalkomputer.com/Bahasan/Teknik-Kompilasi/Topik/Analisis-Leksikal>
- [3] Gunnerson, E, 2001, A Programmer's Introduction to C#.Apress. Berkeley.
- [4] Nugroho, A. 2009. T-Sql Query dan Programming Ms Sql Server. Andi. Yogyakarta.
- [5] Maulana, I., Fadli, M., dan Irawan, M. S. 2006, Pengembangan aplikasi bank. Fasilkom UI. Jakarta.