

Implementasi Algoritma Colussi Pencocokan Kata Berbahasa Inggris Pada Dokumen Microsoft Office Word

Ahmad Ardianto

Program Studi Teknik Informatika, Universitas BUDI DARMA, Medan, Indonesia

Email: ahmadardianto@gmail.com

Abstrak—Kata berbahasa Inggris sering digunakan di berbagai dokumen termasuk salah satunya adalah pada dokumen Microsoft Office Word. Suatu data atau dokumen dapat mengalami kesalahan yaitu pada kata asing yang tidak sesuai dengan makna atau arti yang sebenarnya sehingga dokumen tersebut tidak dapat digunakan. Untuk mendapatkan hasil dokumen yang tepat dan kata asing yang sesuai dengan makna, hal ini dilakukan dengan menggunakan algoritma colossi dengan melakukan pencocokan kata asing tersebut. Algoritma colossi melakukan pergeseran teks dari arah kiri ke arah kanan. Algoritma colossi memiliki dua tahap penyelesaiannya pertama Perbandingan dilakukan dari kiri ke kanan pada teks yang terletak pada posisi yang sama dengan posisi pola dimana nilai dari fungsi kmp_ext sedikit lebih besar dari -1. Posisi-posisi ini dinamakan nohole(s), kedua Membandingkan posisi-posisi yang tersisa (nama lainnya adalah hole(s)) dari arah kiri ke kanan. Dengan algoritma ini diharapkan dapat menyelesaikan masalah kata asing pada dokumen Microsoft office Word. Penelitian ini dibahas agar proses yang dilakukan menerapkan pencocokan kata asing dengan menggunakan algoritma colossi, maka dibangun aplikasi menggunakan Microsoft Visual Basic 2008, dan diharapkan hasil dari aplikasi dapat menghasilkan dokumen yang tepat.

Kata Kunci: Kata Berbahasa Inggris; Algoritma Colussi

Abstract—English words are often used in various documents including one of them is in Microsoft Office Word documents. A data or document may experience errors, namely in foreign words that do not match the true meaning or meaning so that the document cannot be used. To get the correct document results and foreign words that match the meaning, this is done by using the Colossi algorithm by matching these foreign words. The colossi algorithm shifts the text from left to right. The colossi algorithm has two stages of its completion. First Comparisons are made from left to right on text which is located in the same position as the pattern position where the value of the function kmp_ext is slightly greater than -1. These positions are called nohole (s), second Comparing the remaining positions (other names are hole (s)) from the left to the right. With this algorithm it is expected to solve the problem of foreign words in Microsoft office Word documents. , then the application is built using Microsoft Visual Basic 2008, and it is expected that the results of the application can produce the right document.

Keywords: English Words; Colussi Algorithm

1. PENDAHULUAN

Kata berbahasa Inggris digunakan di berbagai dokumen termasuk salah satunya adalah pada dokumen *Microsoft Office Word*. Suatu data atau dokumen dapat mengalami kesalahan yaitu pada kata asing yang tidak sesuai dengan makna atau arti yang sebenarnya sehingga dokumen tersebut tidak dapat digunakan.

Kata berbahasa Inggris pada dokumen *Microsoft Office Word* sangat penting, karena kata berbahasa Inggris sering digunakan. Masalah kata asing di dalam *Microsoft Office Word* sering terjadi. Masalah ini menyebabkan data atau dokumen yang telah dibuat tidak sesuai dengan yang diharapkan. Jika kata asing tidak sesuai maka akan merubah pengertian dari isi dokumen tersebut.

Dalam penelitian ini, melakukan proses pencocokan teks pada dokumen *Word* (Docx) dengan menggunakan algoritma *Colussi*. Algoritma *Colussi* menerapkan prinsip pencarian dengan melakukan pembagian pola menjadi dua himpunan, pada proses pencarian pertama dilakukan dari kiri kekanan, dan jika tidak ditemukan, maka dilakukan proses pencarian kedua yaitu dari kanan ke kiri. *Field* yang ingin dicari dalam proses pencarian dengan memasukan istilah sebagai *input*-annya, kemudian akan mengetahui arti dari istilah tersebut.

Algoritma *Colussi* merupakan suatu pengembangan dari algoritma *Knuth-Morris-Pratt*. Algoritma *Knuth-Morris-Pratt* sendiri adalah algoritma pencocokan string dengan cara memelihara informasi karakter-karakter sebelumnya untuk melakukan jumlah pergeseran yang lebih jauh. Pada algoritma *Colussi* sendiri, himpunan dari posisi pola dibagi menjadi dua sub himpunan terpisah.

Menurut May Aprina Saragih mencocokkan teks *font italic* dengan menggunakan algoritma *Brute Force* akan menempatkan dan mencari semua kemungkinan kata berbahasa Inggris dengan Mencoba setiap posisi *pattern*.

terhadap teks, kemudian dilakukan proses pencocokan setiap karakter dan teks pada posisi tersebut Mencocokkan teks *font italic* untuk kata berbahasa Inggris yang diterapkan didalam dokumen *Microsoft Office Word*[1]. Menurut Raden Prana Algoritma *Colussi* sebagai salah satu turunan dari algoritma *Knuth-Morris-Pratt* dapat digunakan dalam proses DNA *matching* tersebut. Keunggulan algoritma ini adalah kompleksitas waktu yang cukup rendah dan kemampuan menemukan kecocokan yang cukup cepat[2].

2. METODOLOGI PENELITIAN

2.1 Pencocokan Kata

Pencocokan *string* merupakan bagian penting dari sebuah proses pencarian *string* (*string saearching*) dalam sebuah dokumen. Hasil dari pencarian sebuah *string* dalam dokumen tergantung dari teknik atau cara pencocokan *string* (*string searching*) yang digunakan. Pencocokan *string* (*string searching*).

2.2 Algoritma Colussi

Algoritma Colussi merupakan suatu pengembangan dari algoritma *Knuth-Morris-Pratt*. Algoritma *Knuth-Morris-Pratt* sendiri adalah algoritma pencocokan dengan cara memelihara informasi karakter-karakter sebelumnya untuk melakukan jumlah pergeseran yang lebih jauh. Pada algoritma *Colussi* sendiri, himpunan dari posisi pola dibagi menjadi dua sub himpunan terpisah.

2.3 Dokumen

Kata dokumen berasal dari bahasa latin yaitu *docere*, yang berarti mengajar. Pengertian dari kata dokumen ini seringkali digunakan para ahli dalam dua pengertian, yaitu pertama, berarti sumber tertulis bagi informasi sejarah sebagai kebalikan daripada kesaksian lisan, artefak, peninggalan-peninggalan terlukis, dan petilasan-petilasan arkeologis.

3. HASIL DAN PERANCANGAN

Kata dokumen berasal dari bahasa latin yaitu *docere*, yang berarti mengajar. Pengertian dari kata dokumen ini menurut seringkali digunakan para ahli dalam dua pengertian, yaitu pertama, berarti sumber tertulis bagi informasi sejarah sebagai kebalikan daripada kesaksian lisan, artefak, peninggalan-peninggalan terlukis, dan petilasan-petilasan arkeologis. Pengertian kedua, diperuntukan bagi surat-surat resmi dan surat-surat negara seperti surat perjanjian, undang-undang, hibah, konsesi, dan lainnya. Lebih lanjut, Gottschalk menyatakan bahwa dokumen (dokumentasi) dalam pengertiannya yang lebih luas berupa setiap proses pembuktian yang didasarkan atas jenis sumber apapun, baik itu yang bersifat tulisan, lisan, gambaran, atau arkeologis.

3.1 Penerapan Algoritma Colussi

Implementasi algoritma *Colussi* merupakan tahap penyelesaian untuk melakukan pencocokan *string* dalam pencarian kata berbahasa Inggris pada dokumen. Pencocokan *string* menggunakan algoritma *Colussi* adalah seperti berikut ini :

Teks : *METHOD COLUSSI*

Pattern : *COLUSSI*

Tabel 1. PreColussi

I	0	1	2	3	4	5	6	7
x[i]	C	O	L	U	S	S	I	*
hmax[i]	0	1	2	3	4	5	6	7
kmin[i]	0	1	2	3	4	5	6	0
rmin[i]	7	0	0	0	0	0	0	
h[i]	1	2	3	4	5	6	0	
shift[i]	1	2	3	4	5	6	7	

Langkah-langkah pencocokan *string pattern* dengan teks :

1. Langkah Pertama:

i 0 1 2 3 4 5 6 7 8 9 10 11 12 13

T M E T H O D C O L U S S I

P C O L U S S I

h[i] 1

Pencocokan *string* dilakukan berdasarkan nilai *h[i]* yaitu dimulai dari karakter yang terdapat pada *index* Terlihat pada pencocokan O dan E adalah tidak sama, maka dilakukan pergeseran sebanyak nilai pada *Shift[0]* yaitu 1 langkah.

2. Langkah Kedua:

i 0 1 2 3 4 5 6 7 8 9 10 11 12 13

T M E T H O D C O L U S S I

P C O L U S S I

h[i] 1

Terlihat pada pencocokan O dan E adalah tidak sama, maka dilakukan pergeseran sebanyak nilai pada *Shift[0]* yaitu 1 langkah.

3. Langkah Ketiga:

```
i 0 1 2 3 4 5 6 7 8 9 10 11 12 13
T M E T H O D   C O L U S S I
P           C O L U S S I
h[i]         1 2
```

Terlihat pada pencocokan O dan O adalah sama, maka dilakukan pencocokan selanjutnya yaitu L dan D. Terlihat pencocokan L dan D adalah tidak sama maka dilakukan kembali pergeseran sebanyak nilai pada *Shif[1]* yaitu 2 langkah.

4. Langkah Keempat:

```
i 0 1 2 3 4 5 6 7 8 9 10 11 12 13
T M E T H O D   C O L U S S I
P           C O L U S S I
h[i]           1
```

Terlihat pada pencocokan O dan D adalah tidak sama, maka dilakukan pergeseran sebanyak nilai pada *Shif[0]* yaitu 1 langkah.

5. Langkah Kelima:

```
i 0 1 2 3 4 5 6 7 8 9 10 11 12 13
T M E T H O D   C O L U S S I
P           C O L U S S I
h[i]           1
```

Terlihat pada pencocokan O dan Spasi adalah tidak sama, maka dilakukan pergeseran sebanyak nilai pada *Shif[0]* yaitu 1 langkah.

6. Langkah Keenam:

```
i 0 1 2 3 4 5 6 7 8 9 10 11 12 13
T M E T H O D   C O L U S S I
P           C O L U S S I
h[i]           1
```

Terlihat pada pencocokan O dan C adalah tidak sama, maka dilakukan pergeseran sebanyak nilai pada *Shif[0]* yaitu 1 langkah.

7. Langkah Keenam:

```
i 0 1 2 3 4 5 6 7 8 9 10 11 12 13
T M E T H O D   C O L U S S I
P           C O L U S S I
h[i]           0 1 2 3 4 5 6
```

Terlihat pencocokan *string* pada langkah ketujuh seluruh *pattern* memiliki kecocokan dengan teks, maka pergeseran untuk pencocokan *string pattern* dengan *string* teks diberhentikan[4].

4. KESIMPULAN

Berdasarkan pembahasan dan hasil dari penelitian, maka diperoleh kesimpulan bahwa proses pencocokan kata berbahasa Inggris pada dokumen *Microsoft Office Word* menjadi lebih mudah. Hasil pada penelitian ini menunjukkan bahwa Algoritma *Colussi* yang telah diterapkan pada aplikasi pencocokan kata berbahasa Inggris pada dokumen *Microsoft Office Word* sangat efektif.

REFERENCES

- [1] A. M. Saragih, "Implementasi Algoritma Brute Force Dalam Pencocokan Teks Font Italic Untuk Kata Microsoft Office Word," *Pelita Inform. Budi Darma, Vol. IV, Nomor 3, Agustus 2013 ISSN 2301-9425*, pp. 84–87, 2013.
- [2] D. Haryadi, I. T. Bandung, and J. G. Bandung, "Pencocokan String dengan Algoritma Reverse Colussi," 2011.
- [3] D. A. K. Pratt and R. P. Adikhrisna, "Aplikasi Algoritma Colussi Pada Pencocokan D A dan Perbandingannya," no. 13506105, pp. 1–4, 2008.
- [4] S. Dharwiyanti, "Pengantar Unified Modeling Language," pp. 1–13, 2010