

# Penerapan *String Matching* Boyer-Moore pada Pencarian Daftar Pustaka Berformat APA

Okihita Hasiholan Sihaloho<sup>1</sup>  
Program Studi Teknik Informatika  
Sekolah Teknik Elektro dan Informatika  
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia  
<sup>1</sup>okihita@gmail.com

**Abstrak**—daftar pustaka adalah sistem penulisan referensi untuk merujuk sebuah buku yang digunakan dalam referensi. Makalah ini mencoba menerangkan bagaimana algoritma Boyer-Moore dalam pencocokan string dapat diimplementasikan di dalam pencarian daftar pustaka yang ditulis dalam format yang didefinisikan oleh APA.

**Index Terms**—*string matching*, daftar pustaka, boyer-moore, bibliography

## I. PENDAHULUAN

Daftar pustaka mengandung informasi tentang penggunaan sumber yang dirujuk dalam karya tulis, makalah, dan jenis publikasi ilmiah lainnya. Saat ini, ada dua jenis format penulisan daftar pustaka yang umum digunakan, yaitu format yang didefinisikan oleh organisasi American Psychological Association (APA) dan yang didefinisikan oleh Modern Language Association (MLA). Umumnya, format APA digunakan pada penulisan referensi karya ilmiah berkenaan dengan *social sciences*, sedangkan format MLA digunakan pada karya ilmiah yang berhubungan dengan *liberal arts* – termasuk sains dan rekayasa, dan ilmu humaniora.

Format penulisan daftar pustaka buku APA adalah sebagai berikut:

Author, A. A. (Year of publication). *Title of work: Capital letter also for subtitle*. Location: Publisher.

dan contoh penerapannya adalah sebagai berikut:

Calfee, R. C., & Valencia, R. R. (1991). *APA guide to preparing manuscripts for journal publication*. Washington, DC: American Psychological Association.

Sedangkan format penulisan referensi dalam MLA adalah sebagai berikut:

Lastname, Firstname. *Title of Book*. City of Publication: Publisher, Year of Publication. Medium of

Publication.

dan contoh penerapannya adalah sebagai berikut:

Quinn, George. *The learner's dictionary of today's Indonesian*. St Leonards, NSW, Australia: Allen & Unwin, 2001. Print.

Dalam berbagai jenis sistem penerimaan informasi (*information retrieval*) dan editor teks, penting bagi pengguna untuk dengan cepat menemukan sebagian atau seluruh kecocokan pola kata dan frasa dalam teks. Dalam hal ini, yang akan dikaji adalah penggunaan *string matching* dalam membantu pustakawan menemukan referensi yang dirujuk dalam dokumen di dalam basis data pustaka.

## II. ALGORITMA PENCOCOKAN STRING BOYER-MOORE

### A. Deskripsi Algoritma

Algoritma Boyer-Moore adalah algoritma pencocokan string yang paling pragmatis saat ini, dan diimplementasikan dalam berbagai aplikasi dunia nyata. Algoritma inilah diimplementasikan dalam fungsi pencarian teks dari berbagai editor teks seperti Microsoft Word. Algoritma Boyer-Moore akan lebih cepat ‘melompat’ (menggerakkan pola ke bagian kanan string untuk mencari kandidat kecocokan selanjutnya) jika teks pada string menggunakan banyak jenis karakter.

Algoritma Boyer-Moore memulai pencocokan karakter dari kanan ke kiri. Karakter paling kanan pada pola merupakan karakter pertama yang akan dicocokkan dengan teks. Algoritma ini memiliki dua fase, yaitu fase *preprocessing* dan fase pencarian.

Fase *preprocessing* memiliki dua buah aturan untuk menggeser pola (kata kunci yang dicari) ke arah kanan. Kedua fungsi ini disebut penggeseran **Good Suffix Rule** dan penggeseran **Bad Character Rule**. Fungsi penggeseran GSR disimpan ke dalam sebuah tabel bernama *goodsuf* berukuran  $m+1$ . Sedangkan fungsi

penggeseran *bad-character* disimpan ke dalam sebuah tabel bernama *badchar* yang berukuran  $n$ .

Pembentukan tabel *badchar* dan *goodsuf* mempunyai kompleksitas waktu  $O(m+n)$  dan kompleksitas ruang  $O(m+n)$ . Sedangkan kompleksitas waktu untuk fase pencarian adalah  $O(mn)$ . Kasus terbaik untuk algoritma ini mempunyai kompleksitas waktu  $O(n/m)$  sedangkan pada kasus terburuk akan terjadi sebanyak  $3n$  kali perbandingan untuk pencarian dengan pola yang tidak berulang (periodik).

### B. Pseudokode Algoritma

```
procedure badchar(input x: array of
char, input m:integer, input/output
bmBc: array of integer)
```

**Deklarasi**  
i:integer

**Algoritma**  
for (i=0; i<ASIZE; ++i)  
    bmBc[i] <- m

for (i=0; i < m-1; ++i)  
    badchar[x[i]] <- m-i-1

```
procedure suffixes(input x: array of
char, input m: integer, input/output
suff: array of integer)
```

**Deklarasi**  
f, g, i : integer

**Algoritma**  
suff[m-1] <- m;  
g <- m-1  
for (i = m-2; i >= 0; --i)  
    if (i>g and suff[i+m-1-f]<i-g)  
        suff[i] <- suff[i+m-1-f]  
    else {  
        if (i<g)  
            g <- i  
        f = i  
        while(g> 0 and x[g]=x[g+m-1-f])  
            --g  
        suff[i] <- f-g

```
procedure goodsuf(input x: array of
char, input m:integer, input/output
bmGs: array of integer)
```

**Deklarasi**  
i, j : integer  
suff : array [0..XSIZE] of integer

**Algoritma**  
suffixes(x, m, suff)  
for (i = 0; i < m; ++i)  
    bmGs[i] <- m  
    j <- 0  
    for (i = m-1; i >= -1; --i)  
        if (i = -1 or suff[i] = i+1)  
            for (j = 0; j < m-1-i; ++j)  
                if (bmGs[j] = m)  
                    bmGs[j] <- m-1-i  
    for (i = 0; i <= m-2; ++i)  
        goodsuf[m-1-suff[i]] <- m-1-i

```
procedure checkbiblio(input x: array of
char, input m:integer, input y: array of
char, input n: integer)
```

**Deklarasi**  
i, j : integer  
goodsuf : array [0..XSIZE] of integer  
badchar : array [0..ASIZE] of integer

**Algoritma**  
/\* Praproses \*/  
goodsuf(x, m, mygoodsuf)  
badchar(x, m, mybadchar)  
  
/\* Searching \*/  
j = 0  
while (j <= n-m)  
    for (i = m-1; i >= 0 and x[i]= y[i+j];  
        --i)  
        if (i < 0)  
            OUTPUT(j)  
        j <- j + mygoodsuf[0]  
        else  
            j <- j + MAX(mygoodsuf[i],  
                          mybadchar[y[i+j]]-m+1+i)

### III . PENERAPAN ALGORITMA BOYER-MOORE UNTUK DAFTAR PUSTAKA BERFORMAT APA

Seperti yang disebutkan sebelumnya, jika daftar pustaka dianggap sebagai sebuah string, maka pencarian Boyer-Moore berparameter dapat digunakan untuk mencari atribut tertentu dalam kumpulan bibliografi. Untuk kasus ini, yang akan diuji adalah penulisan pustaka untuk buku berformat APA sebagai berikut:

Landler, M. (2007, June 2). *Bush's Greenhouse Gas Plan Throws Europe Off Guard*. ew York Times, p. A7.

#### IV . KESIMPULAN

Kesimpulan dari makalah ini adalah algoritma Boyer-Moore menggunakan dua buah metode preproses yang berbeda untuk menentukan pergeseran maksimum yang mungkin dilakukan, yaitu *bad character* dan *good suffix* sehingga pencocokan string akan lebih efektif dan efisien.

Metode Boyer-Moore ini dapat digunakan sebagai salah satu alternatif untuk pencocokan pola pencarian dalam sebuah string referensi daftar pustaka. Karena daftar pustaka menggunakan banyak karakter, termasuk karakter khusus seperti tanda kurung, titik dua, koma, dan angka untuk penulisan tahun, maka penggunaan Boyer-Moore akan lebih efektif ketimbang algoritma KMP dalam pencarian string dalam basis data referensi.

#### DAFTAR PUSTAKA

- [1] Lang H.W., String Matching Boyer-Moore Algorithm, <http://www.itl.fhflensburg.de/lang/algorithmen/pattern/bmen.htm>, diakses tanggal 19 Desember 2013
- [2] library.flcc.edu/APA\_FLCC.pdf, diakses tanggal 19 Desember 2013
- [3] Munir, Rinaldi, *Diktat Kuliah Strategi Algoritmik*. p193.

#### PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 20 Desember 2013

ttd

Okihita Hasiholan Sihaloho, 13509058