IMPLEMENTASI ALGORITMA KNUTH-MORRIS-PRATT DALAM APLIKASI UNTUK PENERJEMAHAN IDIOM BAHASA INGGRIS

Rahadian Djati Pramono[1)], Selvia Lorena Br.Ginting[2)]
[1),2)]Jurusan Sistem Komputer Universitas Komputer Indonesia, Bandung
rahadiandjatip@gmail.com

ABSTRAK

Bahasa Inggris merupakan bahasa International yang dijadikan sebagai penghubung berbagai macam komunitas di dunia. Meskipun bukan bahasa yang paling banyak digunakan di dunia akan tetapi Bahasa Inggris merupakan bahasa yang wajib dipelajari di setiap Negara. Hal ini memicu para developer software untuk menciptakan sebuah perangkat lunak yang berguna sebagai translator seperti Google Translate dan Bing Translator yang akan menerjemahkan Bahasa Inggris ke berbagai macam bahasa. Akan tetapi ketika menemukan sebuah kata yang berbentuk idiom maka hasil terjemahan tersebut tidak memberikan hasil yang sesuai dengan kata idiom yang ingin di terjemahkan karena akan diterjemahkan kata per kata. Pada penelitian ini akan menggunakan Algoritma string matching KMP (Knuth-Morris-Pratt) yang akan berfungsi sebagai filter yang akan menyimpan informasi dari idiom Bahasa Inggris tersebut.

Kata Kunci: Idiom Bahasa Inggris, Algoritma Knuth-Morris-Pratt, Translator, String Matching

1. PENDAHULUAN

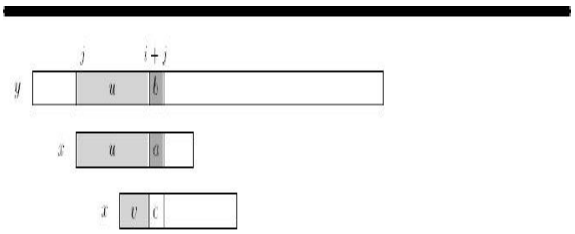
Bahasa merupakan suatu bentuk alat komunikasi manusia yang berupa lambang bunyi melalui alat ucap, dimana setiap suara yang dikeluarkannya memiliki arti untuk berinteraksi dengan manusia lainnya. Setiap bahasa di dunia berbeda-beda untuk menghubungkan perbedaan tersebut, diperlukan suatu bahasa yang dapat digunakan secara global yaitu Bahasa Inggris. Kemajuan teknologi pun berdampak positif pada perkembangan Bahasa Inggris.

Setiap informasi pada teknologi misalnya internet menggunakan Bahasa Inggris sebagai alat komunikasinya. Kurangnya kemampuan berbahasa Inggris merupakan kendala dalam berkomunikasi menggunakan teknologi. Orang Indonesia sering memiliki kendala pada pembendaharaan kata (*vocabulary*) maupun tata bahasa (*grammar*) dalam Bahasa Inggris. Oleh karena itu dibangunlah aplikasi-aplikasi untuk menutupi kendala tersebut, contohnya kamus *online* dan mesin penerjemah seperti pada Google Translate, namun Google Translate seringkali memberikan hasil yang tidak tepat pada saat penerjemahan terutama pada idiom. Idiom merupakan gabungan kata yang memiliki arti yang baru. Sebagai contoh kata "*give*" memiliki arti memberi dan kata "*up*" memiliki arti atas, apabila kedua kata tersebut digabungkan "*give up*" dan diartikan maka hasilnya berbeda dari kata dasarnya yaitu menyerah. Oleh karena itu digunakan algoritma *string matching Knuth-Morris-Pratt* sebagai pencocok *filter* dari idiom tersebut.

2. LANDASAN TEORI

Algoritma *Knuth-Morris-Pratt* adalah salah satu algoritma pencarian *string*, dikembangkan secara terpisah oleh Donald E. Knuth pada tahun 1967 dan James H. Morris bersama Vaughan R. Pratt pada tahun 1966, namun keduanya mempublikasikannya secara bersamaan pada tahun 1977.

Algoritma *Knuth-Morris-Pratt* merupakan salah satu algoritma pencarian *string*, algoritma ini memelihara informasi yang digunakan untuk melakukan jumlah pergeseran pada setiap kali ditemukan ketidak cocokan *pattern* dengan teks. Algoritma ini menggunakan informasi tersebut untuk membuat pergeseran yang lebih jauh, tidak hanya satu karakter. Dengan menggunakan algoritma *Knuth-Morris-Pratt*, waktu pencarian dapat dikurangi secara signifikan.



Gambar 1. Simulasi Kerja Algoritma KMP

Proses dengan menggunakan algoritma KMP akan seperti berikut:
Misalkan *pattern* = "abcabd" akan dicari pada sebuat teks = "abcabcabd", prosesnya adalah sebagai berikut.

a	b	c	a	b	c	a	b	d
a	b	c	a	b	d			

Gambar 2. Proses awal pencocokan *string*

Pencocokan pertama yaitu samakan ujung kiri *pattern* dengan ujung kiri teks. Karakter-karakter pada posisi 1..5 sama (*match*), tetapi karakter c dan d pada posisi 6 menghasilkan ketidakcocokan. Untuk menentukan jumlah pergeseran yang akan dilakukan, mula-mula harus dilakukan perhitungan fungsi pinggiran untuk karakter yang sama atau cocok. Fungsi pinggiran adalah sebagai ukuran awalan terpanjang dari suatu *string* P yang merupakan akhiran dari P[1..j]. Fungsi pinggiran untuk *pattern* "abcabd" akan didefinisikan sebagai berikut:

j	1	2	3	4	5	6
P[j]	a	b	c	a	b	D
B(j)	0	0	0	1	2	0

Gambar 3. Pencarian fungsi pinggiran (pra-pencarian)

Jumlah pergeseran selanjutnya ditentukan oleh fungsi pinggiran dari awalan P yang bersesuaian (*match*). Dalam kasus ini, awalan yang bersesuaian adalah "abcab", yang memiliki panjang 5. Dari fungsi pinggiran diatas dapat dilihat bahwa pinggiran yang terpanjang untuk *string* P[1..5] adalah "ab" yang panjangnya bernilai 2 karakter. Maka berdasarkan fungsi pinggiran diatas, jarak pergeseran adalah 5–2=3. Jadi *pattern* P digeser sejauh 3 karakter dan perbandingan dilakukan mulai dari posisi j = 3 yang dihitung dari awal *pattern*.

a	b	c	a	b	c	a	b	d
			a	b	c	a	b	d

Gambar 4. Hasil akhir pergeseran

Setelah dilakukan pergerseran yang berdasarkan fungsi pinggiran, maka pada akhirnya ditemukan kecocokan dalam string diatas.

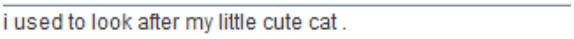
Idiom adalah entitas leksikal yang lebih berfungsi sebagai sebuah kata, walapun terdiri atas beberapa kata (Katamba, 1994:291). Kridalaksana (2007) mendefinisikan idiom sebagai konstruksi yang maknanya tidak sama dengan makna komponennya. Kridalaksana juga membedakan idiom dari semiidiom. Semiidiom menurutnya adalah konstruksi yang salah satu komponennya mengandung makna khas yang ada dalam konstruksi itu saja, misalnya *mata kaki* dan *harga diri*.

Idiom bahasa Inggris adalah suatu ungkapan yang terbentuk dari gabungan beberapa kata bahasa Inggris yang akan menghasilkan arti baru. Dalam idiom kita

tidak dapat mengartikan atau menerjemahkan melalui arti kata pembentuk idiom itu, karena akan membuat kata itu berbeda arti. Jadi intinya idiom itu gabungan kata atau *phrase* yang memiliki satu arti. Seperti pada kata "*looking for*". Sebagai orang yang awam, saat membaca atau pun mendengar kata-kata tersebut, yang terpikirkan mengenai artinya adalah melihat untuk dan jika ditelaah lebih lanjut, makna tersebut tidaklah masuk akal. Berdasarkan kamus idiom kata "*looking for*" memiliki arti mencari.

3. IMPLEMENTASI DAN PENGUJIAN

Data kelas kata idiom mengacu pada kamus idiom Windy Novia. Data juga meliputi kata-kata lainnya yang mengacu pada kamus Echols & Shadily yang di dalamnya meliputi beberapa kelas kata yaitu kata kerja (*verb*), kata benda (*noun*), kata sifat (*adjective*), kata keterangan (*adverb*), kata ganti (*pronoun*), kata depan (*preposition*), kata hubung (*conjunction*), kata seru (*interjection*), kata sandang (*article*), kata ganti kepemilikan (*possesive pronoun*), kata ganti depan. Dalam keseluruhan data tidak terdapat kata benda dalam bentuk jamak (plural) seperti yang telah disebutkan dalam batasan masalah. Pengujian dilakukan dengan menggunakan kalimat "*i used to look after my little cute cat*"



Gambar 5. *Input text* percobaan pertama

Untuk menganalisis kata-kata apa saja yang termasuk dalam kata idiom, diperlukan melakukan pengecekan untuk setiap kata apakah kata-kata tersebut termasuk pada kelas kata idiom pada *database*. Berdasarkan dari sampel pengujian pada Gambar 5, kata-kata yang termasuk pada kelas kata idiom dapat dilihat pada Tabel 1.

Urutan Kata	Termasuk Dalam Kelas Kata Idiom
2	used
3	to
4	look
5	After

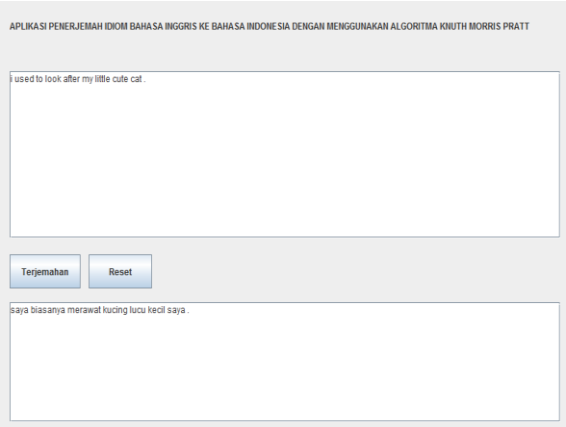
Tabel 1. Tabel pengujian sampel pertama pencocokan kata

Dengan melakukan pencocokan menggunakan algoritma *Knuth-Morris-Pratt* dengan menggunakan kata-kata pada *database* idiom. Kata idiom yang dihasilkan dari pencocokan tersebut dapat di lihat pada Tabel 2 berikut.

No	Hasil Pencocokan Idiom	Kelas Kata	Arti
1	Used to	Idiom	biasanya
2	Look after	Idiom	merawat

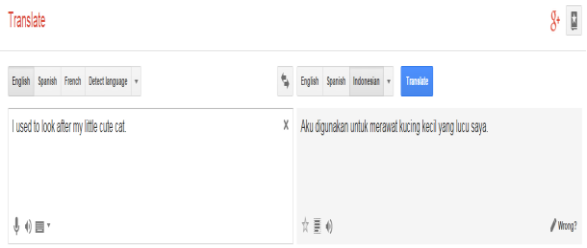
Tabel 2. Tabel pencocokan idiom sampel pertama

Berdasarkan sampel pengujian, *output* yang dihasilkan sesuai dengan analisis dan perancangan yang dilakukan sebelumnya dapat dilihat pada Gambar 6 berikut.



Gambar 6. *Output* terjemahan

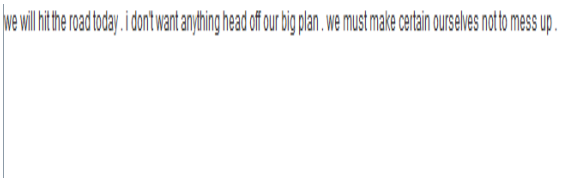
Untuk melakukan perbandingan, dengan *input* yang sama, dilakukanlah penerjemahan menggunakan *Google Translate*. Hasil terjemahan yang diberikan *Google Translate* dapat dilihat pada Gambar 7.



Gambar 7. *Output google translate*

Berdasarkan hasil yang ditunjukkan pada Gambar 6, idiom yang terdapat pada daerah *input* adalah *used to, look after*. Hasil terjemahan yang diberikan *Google Translate* pada Gambar 7 tidak sepenuhnya tepat dalam menerjemah idiom yang terdapat dalam *input* yaitu pada idiom *used to* dimana idiom tersebut masih diterjemahkan per kata.

Ambil sampel kalimat berikutnya untuk dilakukan pengujian. Contoh sampel pengujian berikutnya dapat dilihat pada Gambar 4.6 dan *input*-nya, yaitu: "*we will hit the road today, i don't want anything head off our big plan. We must make certain ourselves not to mess up*"



Gambar 8. *Input text* percobaan kedua

Berdasarkan dari sampel pengujian pada Gambar 8, kata-kata yang termasuk pada kelas kata idiom dapat dilihat pada Tabel 3.

Urutan Kata	Termasuk Dalam Kelas Kata Idiom
3	Hit
4	The
5	Road
11	Head
12	Off
18	Make
19	Certain
22	To
23	Mess
24	Up

Tabel 3. Tabel pengujian sampel kedua pencocokan kata

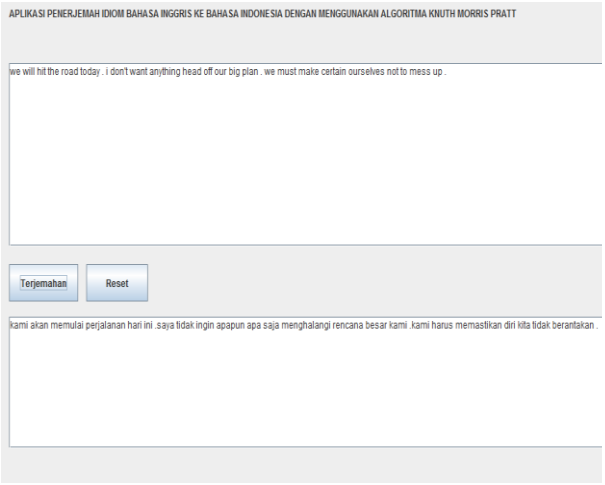
Dengan melakukan pencocokan menggunakan algoritma *Knuth-Morris-Pratt* dengan menggunakan kata-kata pada *database* idiom. Kata idiom yang dihasilkan dari pencocokan tersebut dapat di lihat pada Tabel 4. berikut.

No	Hasil Pencocokan Idiom	Kelas Kata	Arti
1	Hit the road	Idiom	Biasanya
2	Head off	Idiom	menghalangi

3	Make certain	Idiom	Memastikan
4	To mess up	Idiom	berantakan

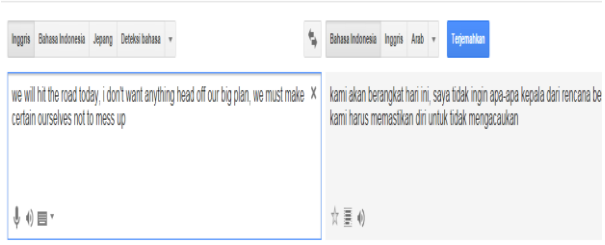
Tabel 4 Tabel pencocokan idiom sampel kedua

Berdasarkan sampel pengujian, *output* yang dihasilkan sesuai dengan analisis dan perancangan yang dilakukan sebelumnya dapat dilihat pada Gambar 9 berikut.



Gambar 9. *Output* terjemahan sampel kedua

Untuk melakukan perbandingan, dengan *input* yang sama, dilakukanlah penerjemahan menggunakan *Google Translate*. Hasil terjemahan yang diberikan *Google Translate* dapat dilihat pada Gambar 10.



Gambar 10. *Output google translate* sampel kedua

Berdasarkan hasil yang ditunjukkan pada Gambar 4.7, idiom yang terdapat pada daerah *input* adalah *hit the road, head off, make certain* dan *to mess up.* Hasil terjemahan yang diberikan Google Translate pada Gambar 4.8 tidak sepenuhnya tepat dalam menerjemah idiom yang terdapat dalam *input* yaitu pada idiom *head off* dimana idiom tersebut masih diterjemahkan per kata.

4. KESIMPULAN

Berdasarkan pembahasan sebelumnya dapat diambil kesimpulan bahwa,

1. Algoritma *Knuth-Morris-Pratt* dapat menerjemahkan idiom bahasa Inggris dengan baik
2. Algoritma *Knuth-Morris-Pratt* juga dapat mengatasi kekeliruan dalam mengartikan idiom bahasa Inggris tersbut.

5. DAFTAR PUSTAKA

[1] Echols, J.M. & Shadily, H. 1996. *Kamus Inggris-Indonesia*. Jakarta: Gramedia Pustaka Utama.

[2] http://informatika.stei.itb.ac.id/~rinaldi.munir/Stmik/2012-2013/Makalah2012/Makalah-IF3051-2012-093.pdf

[3] http://repository.usu.ac.id/handle/123456789/40484

[4] Hariyanto, Bambang. 2014. *Esensi-Esensi Bahasa Pemrograman JAVA Revisi Keempat.* Bandung: Informatika

[5] Djuharie, O. S., 2013, *Teknik dan Panduan Meneerjemahkan Bahasa Inggris-Bahasa Indonesia.* Bandung: Yrama Widya

[6] Novia, Windy. 2009. *Idiom Dictionary English-Indonesian.* Wipress

[7] Ardiansyah, Andre. Tanpa Tahun. *Pocket Dictionary Inggris-Indonesia Indonesia-Inggris.* Surabaya: Pustaka Agung Harapan.

[8] Supardi, Yuniar. Tanpa Tahun. *Pemrograman Database Dengan Java & MYSQL*. Jakarta: PT Elex Media Komputindo