

Aplikasi Pengolah Bahasa Alami Untuk Operasi Query Database

Agus Purwo Handoko¹

Email : kenarok2080@gmail.com

ABSTRAKSI

Aplikasi pengolahan bahasa alami untuk pengoperasian query seleksi dapat : untuk aplikasi yang dibuat dapat mengoperasikan query seleksi dengan menggunakan kalimat bahasa Indonesia sederhana. Dasar perintah operasi query yang digunakan adalah aturan Select Operation, sehingga aplikasi yang dibuat hanya dapat digunakan untuk menampilkan data saja dan sebatas data dari 1 tabel saja dan tata bahasa (grammar) yang mengatur bentuk perintah operasi harus ditentukan dalam bentuk aturan produksi sehingga tidak sembarang kalimat dapat diproses. Sistem ini mampu memberi pesan kesalahan ketika ada ketidaksesuaian bahasa alami yang diberikan, namun sebaiknya dimunculkan kalimat alternatif yang berhubungan dengan kesalahan tersebut.

Kata kunci: bahasa alami, natural language processing, operasi query database, select operation

PENDAHULUAN

Natural Language Processing (NLP) merupakan salah satu aplikasi *Artificial Intelligence (AI)* yang dikembangkan agar komputer mengerti dan memahami bahasa alami yang diberikan dan memberi respon hasil pengolahan sesuai yang diinginkan. Kaplan (1989) membangun aplikasi pengolah bahasa alami untuk menyimpan citra dalam basis data, sedang Harada (1997) mengembangkan pengolah bahasa alami untuk melakukan pencarian citra yang disimpan pada basis data. Maragoudakis dkk. (2004) mengembangkan interaksi manusia dan komputer dengan pengolahan bahasa alami pada aplikasi sistem perawatan medis, yang memudahkan pengguna (dokter, paramedic, mahasiswa) berkomunikasi dengan sistem.

Pengolahan bahasa alami menganalisis input bahasa dengan cara mengidentifikasi sintak, semantik dan konteks yang terkandung dalam satu kalimat agar bisa sampai pada kesimpulan untuk memberikan jawaban. Komputer bisa merespon bahasa yang diberikan oleh user, sesuai dengan yang ditargetkan sistem. Pada sistem ini komputer ditargetkan untuk merespon bahasa alami yang diberikan

1. Dosen STMIK Sinar Nusantara Surakarta
Jl. Saman Hudi 84 - 86 Surakarta Telp 0271 - 716500

user yang sesuai dengan aturan produksi yang diberikan untuk menghasilkan operasi Query datatable.

Perumusan Masalah

Dalam artikel ini dibatasi hanya untuk mentranslate Bahasa SQL kedalam bahasa Indonesia dan hanya untuk kondisi query yang sederhana. Bahasa alami yang digunakan adalah bahasa Indonesia yang sederhana yang memenuhi *grammar* yang sudah ditetapkan. Data table yang akan dioperasikan dibatasi hanya pada table-table yang dibuat dengan menggunakan pengolah database (DBM) berbasis *Xbase* dan berekstensi .dbf seperti : *dBase*, *FoxPro*, *Visual FoxPro* ataupun *Visual dBase*.

Tujuan dan Manfaat

Tujuan dari Penelitian

1. Untuk memudahkan orang dalam merelasikan table tanpa harus tahu bahasa query database.
2. Mengaplikasikan *Natural Language Processing* (NLP) merupakan salah satu aplikasi *Artificial Intelligence* (AI)

Manfaat Penelitian :

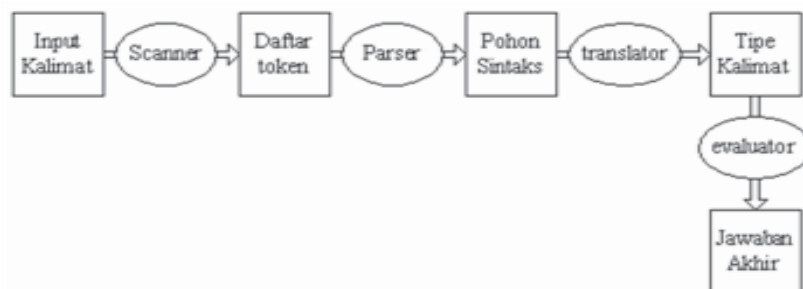
1. Sebagai aplikasi queri database dengan bahasa Indonesia.
2. Bahan referensi bagi peneliti lain yang akan mengembangkan NLP secara lebih detail untuk Query Database secara sempurna yang mungkin sampai *Nested Query*.

Metode Penelitian

Metode penelitian yang digunakan adalah dengan menggunakan metode kepustakaan (membaca referensi) di perpustakaan dan melakukan observasi dengan melakukan surfing di internet.

Pembahasan

Pengolah Bahasa Alami



Gambar 1: Komponen pengolah bahasa alami untuk operasi queri data

Komponen pengolah bahasa alami untuk operasi queri ditunjukkan pada Gambar.1. Setiap kalimat bahasa alami yang dimasukkan akan melewati proses yang dilakukan oleh *scanner*, *parser*, *translator* dan *evaluator* sebelum mendapatkan jawaban akhir. *Scanner* akan melakukan pemeriksaan bentuk kalimat dan mengelompokkannya menjadi daftar *token* yang kemudian diteruskan ke proses berikutnya yang dilakukan oleh *parser*. Dalam proses ini *parser* melakukan pelacakan terhadap *token-token* tersebut untuk dibandingkan dengan daftar *token* yang tersedia dan dicocokkan dengan aturan produksi yang ada. Proses akan dilanjutkan jika aturan produksi dipenuhi. *Translator* akan menterjemahkan hasil parsing untuk menentukan kalimat perintah apa yang diterima. Hasil proses ini akan diteruskan ke *evaluator* untuk melakukan operasi queri.

Bahasa alami yang digunakan pada sistem ini adalah bahasa Indonesia yang mempunyai tata aturan penulisan (*grammar*) tersendiri yang dituliskan dalam bentuk *Backus Naur Form* (BNF) sebagai aturan produksi yang bisa dimengerti oleh sistem. Aturan produksi merupakan suatu kaidah yang menspesifikasikan bagaimana suatu bahasa dibentuk dari transformasi suatu string ke bentuk string lainnya. Bentuk aturan produksi yang dihasilkan dari sebuah kalimat bahasa Indonesia yang dituliskan dalam bentuk BNF adalah <Kalimat> '!' <Subyek> <Predikat> <Obyek>. Simbol <.>. dan '!' adalah simbol metalanguage yang bisa berarti tersusun dari. Elemen yang terdapat dalam tanda < > disebut simbol nonterminal atau *variable*, yang harus didefinisikan dengan aturan BNF yang lain, sedangkan elemen yang lain adalah simbol terminal yang tidak dapat didefinisikan dengan aturan yang lain dan harus merupakan symbol (bagian) dari kalimat yang benar.

Pada dasarnya bahasa alami yang digunakan untuk melakukan operasi-operasi queri mengikuti pola tertentu yang disajikan dengan aturan produksi tertentu (Harada, 1997) dimulai dengan simbol awal <S> sebagai berikut (Irawan,2003) :

```
<S> '!' Tampilkan <Field> dari <Datatable>
<S> '!' Tampilkan <Field > dari <Datatable> berdasar <Kondisi>
<S> '!' Tampilkan <Field > dari <Datatable> berdasarkan <Kondisi>urut
<Pengurutan>
<Datatable> '!' Table dalam database
<Ekstensi> '!' .dbf
<field> '!' (<Field1> (<Field2> |
<Field3>* ) | <Seluruh Field>
<kondisi> '!' Filter data yang akan ditampilkan
<Pengurutan> '!' Field untuk pengurutan data
Arti notasi yang digunakan dalam pembuatan aturan produksi tersebut :
```

'!' = didefinisikan sebagai

<> = simbol nonterminal

| = atau

Beberapa bentuk bahasa alami untuk melakukan operasi queri datatable yang diperoleh dari aturan produksi diatas adalah sebagai berikut :

- Tampilkan Seluruh Field dari mahasiswa
- Tampilkan nim, nama, alamat dari mahasiswa
- Tampilkan nim, nama, alamat dari mahasiswa berdasarkan alamat = 'Yogyakarta'
- Tampilkan nim, nama, alamat dari mahasiswa berdasarkan alamat = 'Yogyakarta' urut nim
- Tampilkan nim, nama, alamat dari mahasiswa berdasarkan alamat = 'Yogyakarta' urut nim
- Tampilkan seluruh field dari mahasiswa urut nama

Operasi Querri Database

Operasi Querri database didasarkan pada aturan dalam relasional aljabar untuk melakukan seleksi data (*select operation*), yaitu memilih tuples yang memenuhi predikat yang diberikan. Bentuk umum dari operasi select adalah :

$$\sigma_P(r)$$

Contoh :

$$\sigma_{nim = '250789'}(mahasiswa)$$

$$\sigma_{alamat = 'Yogyakarta' \wedge minat = 'Cerdas'}(mahasiswa)$$

Seleksi data tanpa kondisi

Operasi seleksi tanpa kondisi merupakan operasi untuk menyeleksi data baik seluruh field maupun field-field tertentu dari sebuah table (datatable) seluruh isi data tanpa dibatasi oleh filter tertentu. Operasi ini digunakan untuk menampilkan isi seluruh data. Bentuk dari operasi ini adalah :

- **SELECT * FROM *tablename***
- **SELECT *field1, field2, field3* FROM *tablename***

Contoh :

- **SELECT * FROM *mahasiswa***
- **SELECT *nim, nama, alamat* FROM *mahasiswa***
- **SELECT * FROM *customer***
- **SELECT *kode, nama, alamat, piutang* FROM *customer***

Seleksi data dengan kondisi

Sama dengan operasi seleksi tanpa kondisi, operasi seleksi data dengan kondisi juga dimaksudkan untuk mengambil data dari table baik seluruh field maupun field-field tertentu saja berdasarkan suatu kondisi (filter) tertentu. Filter ini digunakan untuk membatasi data yang akan ditampilkan dari table tersebut. Bentuk dasar dari operasi seleksi dengan kondisi adalah sebagai berikut :

- **SELECT * FROM *tablename* WHERE *condition***
- **SELECT *field1, field2, field3* FROM *tablename* WHERE *condition***

Contoh hasil dari operasi ini adalah :

- **SELECT *nim, nama, alamat* FROM *mahasiswa* WHERE *alamat* = 'Yogyakarta'**

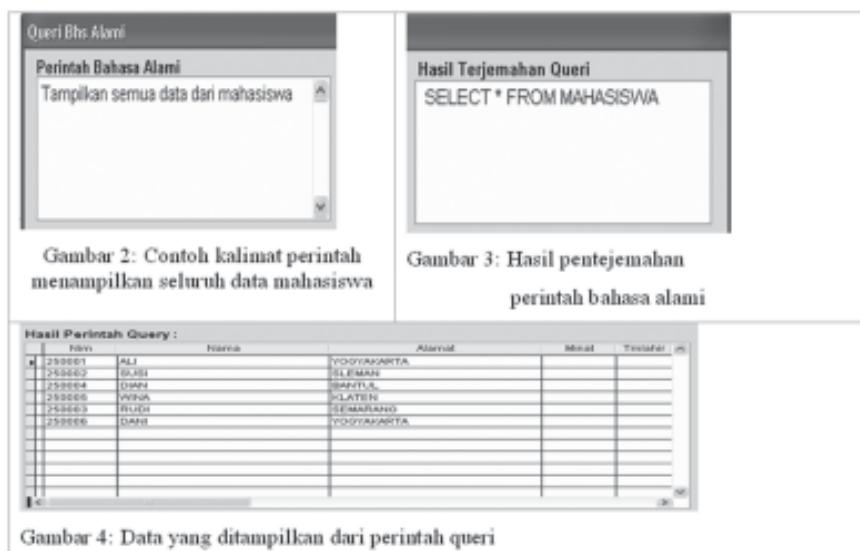
NIM	NAMA	ALAMAT
250001	Ali	Yogyakarta
250006	Wina	Yogyakarta

Hanya record-record yang memenuhi kondisi saja yang ditampilkan, sebagai contoh record data mahasiswa yang beralamat di '**Yogyakarta**' seperti pada table di atas.

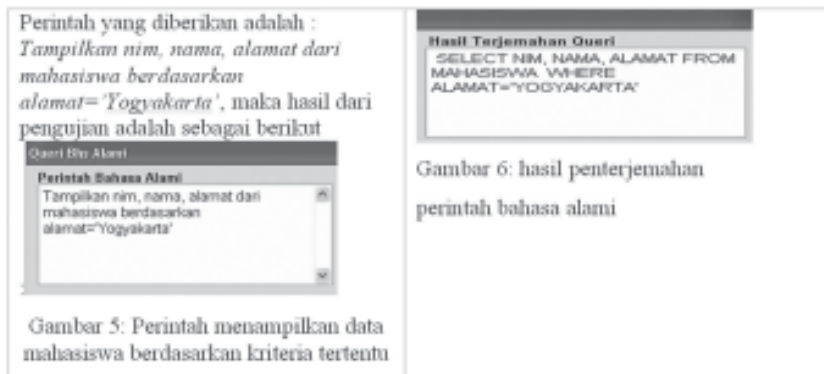
Hasil Uji Coba

Berdasarkan aturan produksi yang telah ditetapkan maka sistem hanya bisa menerima masukan yang sesuai, misalkan kalimat yang dimasukkan adalah '*Tampilkan semua data dari mahasiswa*'. Proses awal yang dilakukan terhadap kalimat tersebut adalah pembentukan daftar *token* yang dilakukan oleh scanner. *Token-token* ini akan diproses oleh parser. Parser melakukan pelacakan terhadap pembentukan kalimat yang kemudian dianalisa kesesuaiannya dengan aturan produksi yang. Penterjemahan kalimat hasil dari pohon sintaks dilakukan oleh *translator* yang menghasilkan tipe kalimat. Dalam proses ini akan diketahui apakah kalimat perintah yang dimasukkan itu sesuai dengan aturan produksi yang ditetapkan atau tidak untuk mendapatkan jawaban akhir yang diinginkan user. Bila sesuai, maka tipe kalimat diproses oleh *evaluator* untuk mendapatkan hasil operasi query seleksi yang diinginkan.

Beberapa contoh hasil pengetesan disajikan untuk melihat beberapa hasil operasi query seleksi yang menunjukkan kalimat perintah untuk menampilkan beberapa perintah query untuk menampilkan beberapa data sesuai dengan kriteria yang diinginkan.



Contoh hasil pengujian lain dari query seleksi dengan bahasa alami adalah sebagai berikut :



Hasil Perintah Query :

Nim	Nama	Alamat
250001	ALI	YOGYAKARTA
250008	DANI	YOGYAKARTA

Gambar 7: hasil eksekusi dari perintah query

Bentuk lainnya apabila pengguna menginginkan pengurutan dari data yang akan ditampilkan. Sebagai contoh perintah yang diberikan adalah : *Tampilkan semua data mahasiswa urut nama*, maka hasil pengujian ditunjukkan sebagai berikut :

Query Bhs Alami

Perintah Bahasa Alami

tampilkan semua data dari mahasiswa urut nama

Hasil Terjemahan Query

```
SELECT * FROM MAHASISWA ORDER BY NAMA
```

Gambar 8: Perintah menampilkan data mahasiswa dan diurutkan berdasarkan nama

Gambar 9: hasil penterjemahan perintah bahasa alami

Hasil Perintah Query :

Nim	Nama	Alamat	Minat	Tinggi
250001	ALI	YOGYAKARTA		
250005	DANI	YOGYAKARTA		
250006	DANI	SEPTUL		
250003	RUDI	DEWAKABO		
250002	SUDI	OLENDA		
250005	WIRI	PLATEN		

Gambar 10: hasil eksekusi dari perintah query

KESIMPULAN

Dari hasil pembuatan aplikasi pengolahan bahasa alami untuk pengoperasian query seleksi dapat disimpulkan sebagai berikut :

1. Aplikasi yang dibuat dapat mengoperasikan query seleksi dengan menggunakan kalimat bahasa Indonesia sederhana.
2. Dasar perintah operasi query yang digunakan adalah aturan Select Operation, sehingga aplikasi yang dibuat hanya dapat digunakan untuk menampilkan data saja dan sebatas data dari 1 tabel saja.
3. Tata bahasa (grammar) yang mengatur bentuk perintah operasi harus ditentukan dalam bentuk aturan produksi sehingga tidak sembarang kalimat dapat diproses.

Sistem ini mampu memberi pesan kesalahan ketika ada ketidak sesuaian bahasa alami yang diberikan, namun sebaiknya dimunculkan kalimat alternatif yang berhubungan dengan kesalahan tersebut.

DAFTAR PUSTAKA

1. Harada. S., Itoh. Y, dan Nakatani. H (1997). "Interactive image retrieval by natural language", *Journal of Optical Engineering*, vol.36, no.12.
2. Hartati, S (1990). "Orientation Adaptive Quadtree Representation", *Tesis Master, Fakultas Ilmu Komputer, University of New Brunswick, Kanada*.
3. Irawan, H.Y (2003). "Penggunaan Bahasa Alami untuk Operasi Antar Image", *Tesis S2 Ilmu Komputer, Universitas Gadjah Mada*.
4. Kaplan M.R (1989). *Constructing Language Processors for Little Languages*. New York: John Wiley & Sons, Inc.
5. Maier, D. dan Warren, D.S (1988) *Computing with Logic. Logic Programming with Prolog*. Menlo Park : The Benjamin/Cummings Publishing Company Inc.
6. Maragaudakis, M., Kladis, B., Tsopanoglou, A., Sgarbas, K., dan Fakotakis, N (2004). "Human-Computer Interaction using Natural Language, An Application in the Medical Treatment Domain", *Technical Report, Wire Communication Laboratory, Department of Electrical Engineering and Computer Technology, University of Patras*
7. Russell, S dan Norvig, P (1995). *"Artificial Intelligence A Modern Approach"*, Prentice Hall, New Jersey.
8. Samet, H. (1994) "The quadtree and related hierarchical data structure", *Computing Surveys*, vol.16,no.2,1984, pp.187-2.