PENGENALAN FRASA BENDA PADA KALIMAT MENGGUNAKAN METODE *SHIFT-REDUCE PARSING*

Novi Yusliani[1], Yunita[2], Wenty Octaviani[3]

[123]*Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Sriwijaya*

[1]novi.yusliani@gmail.com

[2]yunita.v1t4@gmail.com

[3]wentyoct@gmail.com

Abstract— Pada bahasa Indonesia, satuan terkecil linguistik adalah kata. Gabungan dari dua atau lebih kata yang menduduki satu fungsi tersebut dapat membentuk frasa. Frasa sangat berbeda dengan kalimat dikarenakan struktur frasa yang tidak mengandung unsur predikat didalamnya. Pada penelitian ini mengembangkan sebuah perangkat lunak yang dapat mengenali pola frasa benda pada kalimat. Masukan pada perangkat lunak ini adalah kalimat sederhana Bahasa Indonesia. Proses yang dilakukan yaitu, tokenizing, dan Part of Speech (POS) Tagging. Lalu hasil dari proses tersebut akan diolah menggunakan metode Shift Reduce Parsing untuk menemukan pola frasa benda dalam sebuah kalimat tersebut.

Keywords— *Natural Language Processing, Frasa, SRP.*

1. PENDAHULUAN

Sebuah kalimat, akan memungkinkan terdapat frasa didalamnya. Manusia akan dengan mudahnya mengenali frasa tersebut. Akan tetapi komputer tidak bisa mengenali frasa yang terkandung dalam kalimat tersebut. Sehingga tidak bisa membedakan antara kata dan frasa.

Metode yang digunakan untuk mengidentifikasi frasa benda pada kalimat yaitu *Shift Reduce Parsing*. Pada metode ini, ada dua proses yang dilakukan, yaitu *shift* dan *reduce*. Proses shift adalah proses penambahan kata pertama dari RHS ke LHS. Setelah itu, akan dilakukan proses reduce yaitu menghapus elemen dari kata pada LHS dan menggantinya dengan satu elemen non-terminal yang sesuai [1].

2. LANDASAN TEORI

FRASA

Frasa adalah gabungan kata yang mengisi salah satu fungsi sintaksis di dalam kalimat (Mahfudhoh, 2013). Frasa dibagi menjadi beberapa bagian, salah satunya frasa nominal. Contoh frasa benda pada kalimat dapat dilihat pada tabel 1.

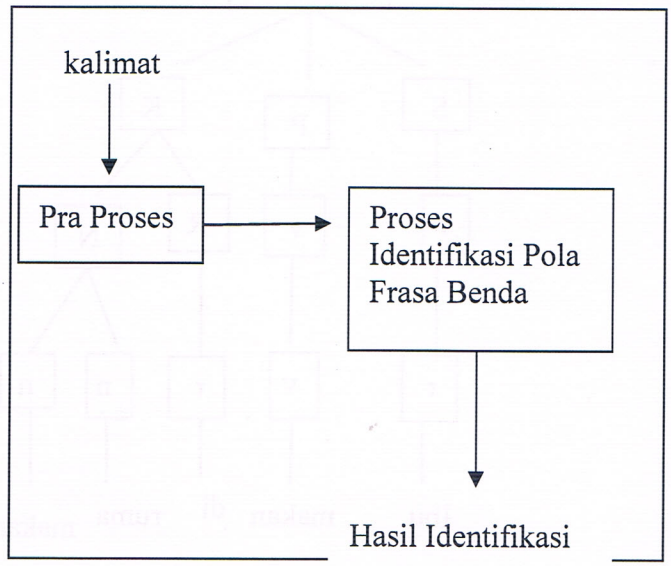
Tabel 1. Contoh frasa benda pada kalimat

No	Kalimat
1	Andi pergi berobat ke rumah sakit setiap minggu
2	Ibu membeli makanan di rumah makan
3	Adik membeli buku baru di toko

Pada kalimat pertama, frasa yang ada pada kalimat tersebut adalah rumah sakit. Rumah sakit adalah gabungan dari rumah dan sakit yang menjelaskan tentang sebuah tempat. Sedangkan pada kalimat kedua rumah makan juga menjelaskan tentang tempat. Frasa pada kalimat ketiga yaitu buku baru yang menjelaskan tentang buku.

METODOLOGI

Perangkat lunak ini menerima masukan kalimat sederhana berbahasa Indonesia. Selanjutnya kalimat tersebut akan di lakukan POSTagger sehingga dihasilkan kata-kata yang menyusun kalimat tersebut. Setelah diketahui kata-kata tersebut, akan ditentukan label kata tersebut. Untuk frasa benda akan didapat dengan menggabungkan pola kata-kata tersebut menggunakan *Shift-Reduce Parsing*. Proses perangkat lunak tersebut dapat dilihat pada bagan berikut ini



Gambar 1. Proses Perangkat Lunak

A. Pra Proses

Pra proses adalah langkah utama untuk mengolah masukan sebelum proses data selesai dilakukan. Pengolahan yang dilakukan yaitu mentransformasikan kalimat menjadi representasi lain. Tujuan proses ini yaitu dilakukan *case folding*, yaitu proses untuk mengubah semua huruf menjadi a-z saja. Selain itu juga menghilangkan tanda baca. Setelah dilakukan proses *case folding,* kalimat tersebut akan dipisah hingga menjadi bentuk per kata. Proses ini disebut tokenizing dengan memisahkan kalimat menggunakan pemisah spasi.

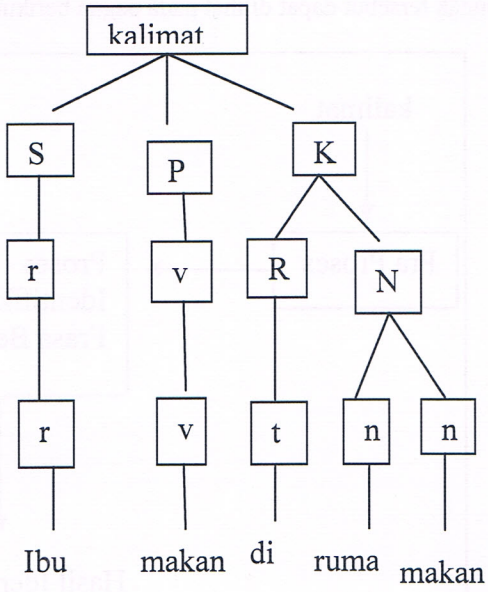
B. Penandaan Kata

Setelah kata pada pra proses telah ditentukan, maka kata-kata tersebut selanjutnya akan diberikan label. Proses ini disebut POST Tagging. Proses pemberian label ini akan menggunakan POS Tagger. Hasil dari POS Tagger adalah kata yang telah diberikan label. Label tersebut menunjukkan fungsi kata di dalam kalimat. Macam-macam pelabelan diantaranya kata benda, kata sifat, kata kerja, dan lain-lain.

POS Tagger ini menggunakan Hidden Markov Model dimana prosesnya memiliki waktu yang lebih baik [3].

C. Pengenalan Pola Frasa Benda

Proses ini yaitu untuk mengenali apakah dalam kalimat masukan tersebut mengandung frasa benda atau tidak. Proses pengenalan ini menggunakan *Shift-Reduce Parsing.* Proses *parsing* menggunakan *Shift Reduce Parsing* dapat dilihat pada gambar 2.



Gambar 2. Pohon Parsing

3. HASIL DAN PEMBAHASAN

SHIFT REDUCE PARSING

Shift – Reduce Parsing adalah teknik parsing yang tergolong *bottom-up.* Model parsing ini digunakan sebagai peruntun token dan membentuk barisan produksi untuk membangun pohon parse. SR Parsing menggunakan tumpukan (stack) untuk menjaga urutan *token.* Secara umum, algoritma *Shift – Reduce Parsing* dilakukan dengan proses berikut:

1. *Shift*, menambahkan satu elemen pada *stack.* Aksi *shift* berupa pemindahan *item* pertama dari RHS (*Right Handle Stack*) ke LHS (*Left Handle Stack*).
2. Reduce, mengurangi elemen teratas pada LHS dan menggantikannya dengan menambahkan satu elemen non-terminal yang sesuai.

Jika LHS kosong, hanya aksi *Shift* yang dapat dilakukan. Jika RHS kosong, hanya aksi *Reduce* yang dapat dilakukan. Jika LHS dan RHS tidak kosong, maka terdapat kemungkinan aksi yang terjadi adalah keduanya, dan pemroses harus memberikan satu kondisi untuk menetukan aksi yang dilakukan. Jika aksi yang dilakukan adalah *Reduce,* maka ditentukan suatu non-terminal apa yang harus ditambah kedalam LHS menggantikan item teratas dari LHS itu sendiri. Jika aksi yang dilakukan *Shift,* maka akan terbentuk suatu *node* terminal baru sebagai *leaf* dari pohon *parsing* dan akan terbentuk *subtree* baru.

IMPLEMENTASI

Implementasi perangkat lunak telah dilakukan dengan memasukan kalimat pada formMain. Jumlah kalimat yang mengandung frasa benda pada percobaan ini ada sebanyak 20 kalimat. Dari hasil uji coba yang dilakukan, ditemukan 3 kalimat yang frasa bendanya tidak terdefinisi. Kesalahan ini dikarenakan pola frasa benda pada program ini dibuat hanya untuk frasa yang terdiri dari dua kata saja, sedangkan untuk yang terdiri lebih dari dua tidak ditangani.

4. KESIMPULAN

Kesimpulan yang didapat pada penelitian ini yaitu : *Shift – Reduce Parsing* dapat digunakan untuk mengidentifikasi pola frasa benda pada kalimat. Dari kesalahan yang ditemukan saat pengujian, diperlukan lagi pengetahuan mengenai pola frasa benda untuk yang terdiri lebih dari dua kata.

ISSN: 2461-0739

DAFTAR REFERENSI

[1] Fachrurrozi, M., Yusliani, N., & Agustin, M. M. (2014). Identification of Ambigous Sentece Pattern in Indonesian Using Shift-Reduce Parsing. 1.

[2] Mahfudhoh, A. L. (2013). Perbedaan struktur frasa nomina dalam penerjemahan Indonesia-Inggris puisi Hulahop bianglala dan rainbow hulahop. ISSN.

[3] Wicaksono, A. F., & Purwarianti, A. (2010). HMM Based Part-of-Speech Tagger for Bahasa Indonesia. Proceeding the Forth International MALINDO Workshop.

Biodata Penulis

Novi Yusliani, saat ini adalah Dosen di Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Sriwijaya Palembang.