

Penerapan Metode *Left Corner Parsing* Dan Analisis Kontekstual Pada *Natural Language Processing*

Deny Jollyta^{*1}, Muhammad Zarlis², Gusrianty³, Yulvia Nora Marlim⁴

^{1,3,4}Sekolah Tinggi Ilmu Komputer Pelita Indonesia; Jalan Jendral Ahmad Yani No. 78-88 Pekanbaru

²Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Sumatera Utara, Medan
e-mail: ^{*1}deny.jollyta@lecturer.pelitaindonesia.ac.id, ²m.zarlis@yahoo.com,

³gusrianty@lecturer.pelitaindonesia.ac.id, ⁴yulvia.nora@lecturer.pelitaindonesia.ac.id

Abstrak

Penerjemahan bahasa secara alami sangat membantu dalam memahami setiap kata maupun kalimat berdasarkan arti yang dimaksud. Ragam bahasa yang tidak mudah untuk dimengerti, menimbulkan kesulitan komunikasi. *Natural Language Processing (NLP)* menjadi pilihan untuk menterjemahkan bahasa manusia dengan komputer secara alami. Sebagai salah satu metode parsing, *Left Corner Parsing* menguraikan bahasa mulai dari constituent terbesar hingga terkecil, yakni kata. Penguraian tersebut diikuti dengan merepresentasikan arti atau dikenal dengan interpretasi kontekstual. Studi ini dilakukan untuk menguraikan bahasa Mandailing menggunakan *Left Corner Parsing* yang dilengkapi dengan analisis menentukan maksud dari penggunaan kalimat tersebut. Terdapat 30 kalimat yang diujikan menggunakan *Left Corner Parsing*. Hasil pengujian memperlihatkan bahwa kesesuaian penguraian struktur kalimat dan ketepatan analisis kontekstual terhadap hasil terjemahan mencapai 93,4%.

Kata kunci— *natural language processing, left corner parsing, interpretasi kontekstual*

Abstract

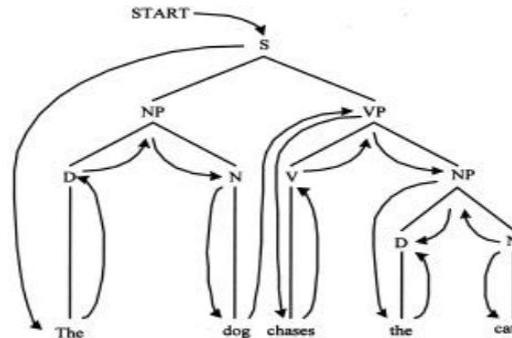
Natural language translation is very helpful in understanding each word or sentence based on the intended meaning. Variety of languages that are not easy to understand, cause miscommunication. Natural Language Processing (NLP) is the choice to translate human language with computer naturally. As one method of parsing, the left corner parsing describes languages ranging from the largest to the smallest constituent, namely the word. The decomposition is followed by representing the meaning or known as contextual interpretation. This study aims to describe the language using the left corner parsing which is equipped with analysis to determine the purpose of sentence using. There are 30 sentences that are tested using Left Corner Parsing. The test results show that the suitability of the decomposition of sentence structure and the accuracy of contextual analysis of the translation results reaches 93,4%.

Keywords— *natural language processing, left corner parsing, contextual interpretation*

1. PENDAHULUAN

Salah satu bidang ilmu dalam *Artificial Intelligence* adalah *Natural Language Processing (NLP)* yang dikembangkan oleh Alan Turing pada tahun 1950 [1]. Penerjemahan berbagai bahasa telah menjadi hal penting untuk menyelesaikan masalah komunikasi antar daerah hingga

antar negara. Persoalan pengolahan bahasa yang paling mengemuka adalah ambiguitas [2], dan ketidaksesuaian bahasa dengan struktur maupun aturan suatu bahasa. Untuk mengatasinya, NLP membahas dan mencari penyelesaian ambiguitas melalui beberapa proses, seperti *scope ambiguity*, *attachment ambiguity*, dan sebagainya [3]. Hal ini memacu perkembangan metode penerjemahan seperti *Left Corner Parsing* yang merupakan gabungan dari *bottom-up* dan *top-down parsing* [4] dengan alur seperti gambar 1 [1], yang dimulai secara *bottom-up* dan diakhiri secara *top-down*.



Gambar 1. Cara Kerja *Left Corner Parsing*

Bahasa Mandailing merupakan salah satu bahasa daerah dari Sumatera Utara. Struktur bahasanya hampir sama dengan struktur bahasa Indonesia. Namun penggunaan kata dalam bahasa Mandailing terkadang menghasilkan struktur berbeda dalam bahasa Indonesia sehingga maksud kalimat juga berbeda. Hal ini mengakibatkan ambigu dalam memahami maksud kalimat. Oleh karena itu perlu dilakukan penelusuran menggunakan *Left Corner Parsing* untuk melihat kesamaan struktur dan kesesuaian maksud kalimat.

Berbagai penelitian tentang penerapan *Left Corner Parsing* telah banyak dilakukan. Pada beberapa penelitian, metode ini diterapkan pada aplikasi penerjemahan bahasa daerah di wilayah Indonesia, seperti bahasa Indonesia ke bahasa Bima [5] dan bahasa Jawa ke bahasa Lampung [6]. Kemampuan *Left Corner Parsing* juga teruji pada keberhasilan dalam menerjemahkan bahasa Inggris [4], [7]. Penelitian [8] menghasilkan sebuah model pendekatan *algorithmic-level* secara paralel dalam pemrosesan kalimat melalui *Left Corner Parsing*.

Untuk memudahkan membaca hasil penerjemahan, dilakukan pengolahan terjemahan dengan beberapa proses, yakni analisis sintaksis, interpretasi semantik dan interpretasi kontekstual, seperti penelitian [1]. Namun dalam penelitian tersebut, penerjemahan hanya sampai pada pembentukan struktur bahasa dengan terjemahan yang tidak mempertimbangkan maksud kalimat sebenarnya. Pada studi ini, selain penelusuran bahasa Mandailing yang menggunakan *Left Corner Parsing*, juga terdapat kontribusi dalam hasil terjemahan kalimat menggunakan interpretasi kontekstual agar hasil terjemahan bahasa Indonesia sesuai dengan yang dimaksudkan dalam bahasa Mandailing.

2. METODE PENELITIAN

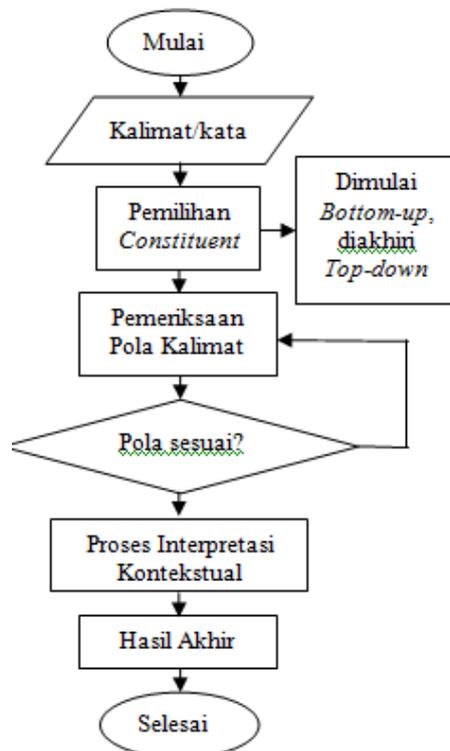
Untuk memudahkan dalam melakukan penerjemahan, dijelaskan bagan alir kerangka konseptual yang mendukung studi ini. Hasil penelusuran penerjemahan bahasa dilakukan dengan beberapa tahapan. Proses penerjemahan sangat tergantung pada pola kalimat dari bahasa yang diterjemahkan. Salah satu tahapan adalah mengetahui kesesuaian pola. Tabel 1 berikut adalah struktur bahasa Indonesia dan bahasa Mandailing yang digunakan dalam penelusuran.

Tabel 1. Struktur Bahasa dan Contoh

Aturan	Bahasa Indonesia	Bahasa Mandailing
S P	S P	S P
<KB><KK>	Ibu mencuci	Umak mambaso
S P O	S P O	S P O
<KB><KK><KB>	Kakak makan sayur	Akak mangan ikkayu
S P K	S P K	S P K
<KB><KK><Ket>	Ayah pergi kerja	Amang kehe karajo
S P O K	S P O K	S P O K
<KB><KK><KB><Ket>	Paman membeli baju kemarin	Tulang manabusi baju natuari

Struktur bahasa dalam tabel 1 memiliki penjelasan, yakni S adalah subjek; P adalah predikat; O adalah objek; KB merupakan kata benda; KG adalah kata ganti; KK adalah kata kerja; dan Ket merupakan keterangan. Struktur ini menjadi acuan dalam melakukan parsing sehingga studi yang dilakukan mengikuti tahapan yang terdapat pada gambar 3.

Langkah kerja dimulai dari membaca bahasa dalam bentuk kata atau kalimat. Studi ini menyediakan 30 kalimat bahasa Indonesia yang diterjemahkan ke bahasa Mandailing. Selanjutnya dilakukan pemilihan *constituent* mulai dari yang terbesar hingga yang terkecil. *Constituent* merupakan bagian yang membentuk kalimat dan berdiri sendiri, seperti kata benda (*noun phrase*), kata kerja (*verb phrase*), dan sebagainya. Pola yang sesuai dengan kaidah tujuan, kemudian dianalisis untuk mendapatkan maksud dari kalimat tersebut dengan menggunakan pola interpretasi kontekstual. Tabel 1 adalah struktur bahasa Indonesia yang telah ditentukan. Penerjemahan bahasa harus memperhatikan pola tersebut dan membandingkannya dengan struktur bahasa Mandailing. Tujuan kesesuaian pola adalah untuk memudahkan parsing dan proses penerjemahan.



Gambar 3. Bagan Alir Kerangka Konseptual

3. HASIL DAN PEMBAHASAN

3.1 *Analisis Data*

Pengujian dilakukan pada data kata atau bahasa yang berasal dari kamus bahasa Indonesia yang kemudian diterjemahkan ke bahasa Mandailing dengan penjelasan data berikut:

- a. Kosakata diambil dari kata dasar bahasa Indonesia dan bahasa Mandailing serta struktur bahasa Mandailing [9], [10].
- b. Data untuk jenis kata diambil dari bahasa Indonesia.
- c. Struktur atau pola bahasa Mandailing mirip dengan bahasa Indonesia. Sebuah struktur kalimat memiliki beberapa yang dituangkan ke dalam *context free grammar* (CFG) yang mempunyai 4-tuple [5], yakni $G = (V, T, P, S)$ {*Grammar/ Tata Bahasa*}.

1) V: {Kal, S, P, K, O1, O2, Pel, KB, KG, KK, KS, Bil, Ket...}

2) T: {gajah, tidur, sakit, kantor, satu, dua...}

3) S : simbol permulaan\

4) {

Kal : <S> <P>

Kal : <S> <P> <O>

Kal : <S> <P> <Pel>

Kal : <S> <P> <K>

Kal : <S> <P> <O> <K>

Kal : <S> <P> <O1> <O2>

Kal : <S> <P> <O1> <O2> <K>

<S> : <KB>|<KG>

<P> : <KB>|<KG>|<KK>|<Bil>

<O> : <KB>|<KG>

<Pel> : <KB>|<KG>|<Bil>|<KK>

<K> : <Ket>

<KB> : {gajah, perut, dan lain-lain}

<KG> : {saya, kamu, dan lain-lain}

<KS> : {sakit, dan lain-lain}

<KK> : {tidur, makan, dan lain-lain}

<Bil> : {satu, dua, dan lain-lain}

<Ket> : {kantor, sekolah, kantor, dan lain-lain}

Analisis data ini berguna untuk menghubungkan pola bahasa pada tabel 1 dengan *constituent* yang dihasilkan dari terjemahan.

3.2 *Penerapan Left Corner Parsing*

Bahasa Indonesia yang diterjemahkan ke dalam bahasa Mandailing berpedoman pada struktur bahasa Indonesia. Penerjemahan bahasa Indonesia ke bahasa Mandailing diuraikan melalui contoh berikut ini.

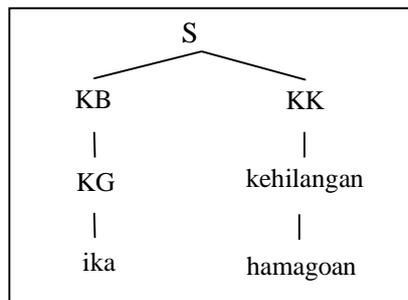
Tabel 2. Data Uji 1

S	P
Ika	Kehilangan

Penelusuran data uji yang terdapat pada tabel 2 menggunakan metode *Left Corner Parsing* adalah sebagai berikut.

- S <KB>|<KK>
 KB <KG>
 KK <KK>
 KG ika
 KS kehilangan
- 1) Masukkan kalimat “ika kehilangan” – mengenali S “ika kehilangan” = S
 - 2) Masukkan kata pertama – ika (KG)
- S
 KG
- 3) KB KG
- S
 |
 KB
 |
 KG
 Ika | kehilangan

Untuk pohon *parsing* (*parse tree*), digambarkan seperti gambar 4.



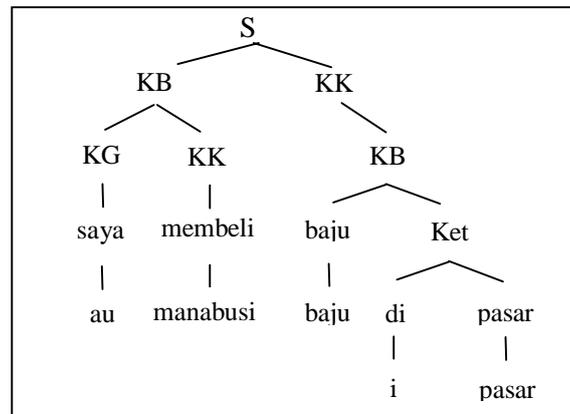
Gambar 4. Pohon *Parsing*

Selanjutnya diberikan contoh kalimat untuk struktur S-P-O-K sebagai berikut.

Tabel 3. Data Uji 2

S	P	O	K
saya	membeli	baju	di pasar

Pada tabel 3, kalimat uji memiliki struktur bahasa Indonesia yang lengkap. Penelusuran dengan metode *Left Corner Parsing* diawali dengan kata pertama “saya”. Pola kata pertama diproses melalui *database* dan mendapatkan bahwa kata “saya” merupakan kata ganti (KG) dari klausa subjek.



Gambar 5. Pohon Parsing Struktur S-P-O-K

Gambar 5 menggambarkan penelusuran yang dimaksud berdasarkan data uji pada tabel 3. Kata kedua yakni “membeli”, merupakan pola kata kerja yang penelusurannya dengan memeriksa kata dalam database dengan susunan pola kata pertama dari KG yang diikuti dengan pola kata kerja (KK). Jika susunan pola sesuai dengan struktur awal, maka susunan kata telah sesuai.

Kata ketiga adalah “baju”. Penelusuran yang sama dilakukan dan menemukan bahwa kata “baju” merupakan KB. Penelusuran terakhir adalah kata “di pasar” yang juga merupakan cabang KB dari kata “baju” dan membentuk pola Ket. Namun kata “di pasar”, terdiri dari dua kata terpisah, yakni “di” dan “pasar”. Hasil akhir keseluruhan dari berbagai kemungkinan dari awal penelusuran menjadi KG-KK-KB-Ket. Format susunan kemudian diuji, disesuaikan dengan pola kata pertama, kedua, ketiga dan keempat terhadap pola struktur awal. Hasil pengujian menunjukkan bahwa susunan kata telah sesuai dengan struktur S-P-O-K.

3.3 Analisis Interpretasi Kontekstual

Proses pengolahan bahasa alami yang telah diterjemahkan, bertujuan untuk memudahkan pengguna mengetahui arti kalimat secara *context-dependent* dan menentukan maksud dari penggunaan kalimat. Secara pola atau struktur kalimat, kalimat “ika hamagoan” dan “au manabusi baju i pasar” telah memenuhi dan memiliki arti yang logis dan mudah dimengerti. Namun ternyata, untuk analisis interpretasi kontekstual, kalimat pertama dan kalimat kedua, dapat memiliki maksud yang berbeda dari awalnya.

Misalnya kalimat “ika hamagoan”, yang berarti “ika kehilangan”. Secara pola dan logika, kalimat ini memenuhi pola KB-KK (S-P) dan dapat diartikan bahwa ika kehilangan sesuatu. Namun jika kata dibalik, maka pola yang terbentuk tidak sesuai lagi (P-S). Kalimat “ika hamagoan” menjadi “hamagoan ika” sehingga terjemahan kalimat dalam bahasa Indonesia adalah “kehilangan ika”. Pada bahasa Mandailing, “kehilangan ika” memiliki pengertian kuat bahwa “ika sudah meninggal dunia”, sedangkan dalam bahasa Indonesia, kalimat yang sama dapat menunjukkan kehilangan sesuatu, kematian ataupun penderitaan.

3.4 Pengujian Hasil Penelusuran

Pengujian kebenaran hasil penelusuran *Left Corner Parsing* dilakukan secara sederhana melalui angket yang disebarkan kepada 30 responden yang merupakan orang Mandailing yang bisa berbahasa Mandailing. Hasil pengisian angket dirangkum dalam Tabel 4.

Tabel 4. Hasil Angket

Kalimat	Jawaban Responden		Persentase
	Sesuai	Tidak Sesuai	
1	28	2	93,3%
2	28	2	93,3%
3	28	2	93,3%
...
19	28	2	93,3%
...
28	28	2	93,3%
29	28	2	...
30	29	1	96,7%
Persentase	93,4%	6,6%	

Angket berisikan ketiga puluh kalimat dalam bahasa Indonesia dan telah diterjemahkan ke dalam bahasa Mandailing. Berdasarkan Tabel 4, terdapat 29 kalimat yang dianggap sesuai hasil terjemahannya oleh 29 responden dan 1 kalimat dianggap tidak sesuai oleh 1 orang responden. Secara keseluruhan, kesesuaian hasil terjemahan mencapai 93,4%.

4. KESIMPULAN

Penerjemahan kalimat atau kata bahasa Indonesia ke bahasa Mandailing secara alami menggunakan metode *Left Corner Parsing*, sangat membantu dalam komunikasi. Studi yang telah dilakukan menunjukkan kesesuaian penelusuran dengan struktur bahasa awal mencapai 93,4%. Hasil interpretasi kontekstual terhadap hasil terjemahan bahasa Mandailing, dianggap tepat karena dapat dimengerti. Adanya analisis bahasa secara kontekstual, memperjelas tujuan kalimat tersebut. Hal ini menjadi penting mengingat perbedaan bahasa daerah yang ada di Indonesia menggambarkan kondisi dan budaya daerah tersebut sehingga sangat penting untuk mengetahui untuk apa dan kapan sebuah kalimat digunakan.

Pada penelitian [1], analisis interpretasi kontekstual tidak ditampilkan. Diharapkan hasil studi ini mampu memberikan informasi dari sudut pandang yang berbeda. Metode *Left Corner Parsing* dengan analisis interpretasi kontekstual memang mudah diaplikasikan. Namun perlu dipahami bahwa harus teliti dalam menyesuaikan dengan kompleksitas kosa kata maupun grammer dari bahasa yang diterjemahkan agar hasil terjemahan lebih tepat dan benar.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Sekolah Tinggi Ilmu Komputer Pelita Indonesia telah mendukung keberhasilan studi ini.

DAFTAR PUSTAKA

- [1] J. Suciadi, 2001, “*Studi Analisis Metode-Metode Parsing dan Interpretasi Semantik pada Natural Language Processing*,” in *Jurnal Informatika*, Vol. 2, No. 1, pp. 13–22, Petra Christian University.
- [2] M. K. Anjali and B. A. P, 2014, “*Ambiguities in Natural Language Processing*,” *Int. J. Innov. Res. Comput. Commun. Eng.*, Vol. 2, No. 5, pp. 392–394.
- [3] C. Bathulapalli, D. Desai, and M. Kanhere, 2016, “*Use of Sanskrit for Natural Language Processing*,” *Int. J. Sanskrit Res.*, Vol. 2, No. 6, pp. 78–81.
- [4] H. Noji, Y. Miyao, and M. Johnson, 2016, “*Using Left-Corner Parsing to Encode Universal Structural Constraints in Grammar Induction*,” in *EMNLP 2016 - Conference on Empirical Methods in Natural Language Processing, Proceedings*, pp. 33–43, doi: 10.18653/v1/d16-1004.
- [5] E. Istiqamah and D. Soyusiawaty, 2017, “*Penerapan Metode Left Corner Parsing Dalam Aplikasi Terjemahan Bahasa Indonesia ke Bahasa Bima*,” in *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, pp. 1–8.
- [6] K. Zuhri and Y. Yuniarthe, 2017, “*Algoritma Left Corner Parsing (LCP) untuk Menerjemahkan Bahasa Jawa ke Bahasa Lampung*,” in *Prosiding Semnas IIB Darmajaya*, pp. 398–410.
- [7] P. Resnik, 1992, “*Left-Corner Parsing and Psychological Plausibility*,” in *Proceedings of Coling*, pp. 191–197, doi: 10.3115/992066.992098.
- [8] N. E. Rasmussen and W. Schuler, 2018, “*Left-Corner Parsing With Distributed Associative Memory Produces Surprisal and Locality Effects*,” *Cogn. Sci.*, Vol. 42, No. 4, pp. 1009–1042, doi: 10.1111/cogs.12511.
- [9] S. Z. Harahap, 2014. “*Reduplikasi Dalam Bahasa Angkola Mandailing*,” *J. Aksara*, Vol. 26, No. 9, pp. 145–156,
- [10] N. Hasibuan, 2014 “*Proses Tipologi Bahasa Mandailing*,” *J. Telangkai Bhs. dan Sastra*, Vol. 8, No. 2, pp. 110–119, , doi: 10.1017/CBO9781107415324.004.