

# Mengatasi Kalimat Ambigu dalam Bahasa Jepang Menggunakan Pohon Sintaks

Muhammad Iqbal Al Khowarizmi (13515086)

*Program Studi Teknik Informatika*

*Sekolah Teknik Elektro dan Informatika*

*Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia*

[13515086@std.stei.itb.ac.id](mailto:13515086@std.stei.itb.ac.id); [iqbalkhowarizmi@gmail.com](mailto:iqbalkhowarizmi@gmail.com)

**Abstrak**—Bahasa Jepang adalah bahasa yang memiliki karakter yang menarik. Kalimat dalam bahasa Jepang pada umumnya memiliki pola SOP atau dengan kata lain, predikat pada umumnya muncul di bagian akhir kalimat. Selain predikat, urutan kata dalam bahasa Jepang tidak terlalu memengaruhi makna. Fleksibilitas seperti ini seringkali memunculkan kalimat-kalimat ambigu. Belum lagi sistem penulisan yang sangat berbeda dan tidak adanya jarak antarkata pada ragam bahasa tulisnya membuat kalimat dalam bahasa Jepang cenderung sulit untuk ditentukan makna semantiknya. Untuk mengatasi kalimat ambigu dan menentukan makna semantiknya, kalimat dalam bahasa Jepang dapat diuraikan menggunakan modifikasi pohon sintaks berbasis ketergantungan.

**Keywords**—ambigu, bahasa Jepang, pohon sintaks, semantik.

## I. PENDAHULUAN

Bahasa Jepang merupakan salah satu bahasa yang paling banyak digunakan saat ini. Pengguna bahasa Jepang pada tahun 2010 diperkirakan mencapai 125 juta pribumi. Universitas di berbagai negara juga menyediakan kuliah bahasa Jepang dan sejumlah sekolah di dunia juga menawarkan pembelajaran bahasa Jepang. Ketertarikan dunia internasional dengan bahasa Jepang juga terus meningkat seiring dengan menyebarnya kebudayaan populer Jepang seperti *anime* dan *manga*.

Sebagai bahasa, bahasa Jepang bisa dikatakan sebagai bahasa yang kontekstual. Artinya, makna kalimat dalam bahasa Jepang sangat mengandalkan konteks di luar makna harafiah kalimat tersebut. Selain itu, dalam kalimat bahasa Jepang, subjek dan objek lazim dihilangkan.

Dalam sistem penulisan resminya, bahasa Jepang menggunakan aksara Tiongkok (*kanji*) dan *kana*, yang asing di mata orang-orang pengguna bahasa dengan aksara latin. Selain itu, kata-kata dalam bahasa Jepang juga tidak dipisahkan oleh spasi seperti dalam bahasa Indonesia atau bahasa Inggris. Sistem penulisan yang

sama sekali berbeda itulah yang menjadikan membaca tulisan dalam bahasa Jepang suatu tantangan tersendiri bagi pelajar bahasa Jepang nonpribumi.

Kedua hal di atas dianggap sebagai kesulitan utama dalam memahami kalimat bahasa Jepang. Akan tetapi, ada cara-cara yang dapat dilakukan untuk memudahkan pemaknaan kalimat dalam bahasa Jepang. Salah satu cara adalah penguraian kalimat menggunakan pohon sintaks. Bagaimana pohon sintaks dapat membantu pemaknaan kalimat dalam bahasa Jepang akan dibahas di dalam makalah ini.

## II. BAHASA JEPANG

### A. Sistem Penulisan

Bahasa Jepang memiliki dua aksara utama yang disebut *hiragana* dan *katakana*. *Hiragana* dan *katakana* merupakan dua versi berbeda dari sekelompok “huruf” yang mewakili bunyi yang sama. Setiap bunyi di dalam bahasa Jepang memiliki representasi “huruf” *hiragana* dan *katakana*-nya. Kedua aksara tersebut masing-masing terdiri dari kurang lebih 50 “huruf”.

Selain kedua aksara tersebut, aksara Tiongkok, yang dalam bahasa Jepang disebut sebagai *kanji*, juga digunakan dalam sistem penulisan bahasa Jepang. Kebanyakan kata dalam bahasa Jepang ditulis menggunakan huruf *kanji* ini. Ada lebih dari 40.000 huruf *kanji*. Sekitar dua ribu di antaranya mewakili 95% karakter yang benar-benar digunakan dalam tulisan bahasa Jepang. Tidak ada jarak (spasi) antarkata di dalam bahasa Jepang tulis sehingga *kanji* sangat diperlukan untuk membedakan kata-kata di dalam suatu kalimat. Terbatasnya bunyi berbeda yang ada di dalam bahasa Jepang membuat homofon lazim ditemukan. *Kanji* juga berfungsi sebagai pembeda untuk homofon-homofon tersebut.

*Hiragana* pada umumnya digunakan sebagai penanda fungsi gramatikal. Terlihat jelas pada penulisan partikel yang semuanya menggunakan *hiragana*. Kata-kata dengan *kanji* yang sulit atau jarang, ekspresi percakapan, dan onomatopoe juga ditulis menggunakan

hiragana.

Meskipun *katakana* merepresentasikan bunyi-bunyi yang sama dengan *hiragana*, *katakana* biasanya digunakan untuk menuliskan kata-kata serapan.



Gambar 1 Hiragana, Katakana dan Kanji.

(sumber: [http://www.workartz.com/wp-content/uploads/2009/10/hiragana\\_katakana\\_kanji.jpg](http://www.workartz.com/wp-content/uploads/2009/10/hiragana_katakana_kanji.jpg))

## B. Penggolongan Kata

Kata-kata dalam bahasa Jepang bisa digolongkan ke dalam setidaknya dua belas kategori. Beberapa dari kategori-kategori tersebut bahkan memiliki beberapa subkategori. Akan tetapi, untuk kemudahan, kategori yang akan digunakan di dalam makalah ini adalah sebagai berikut.

### Nomina

Nomina atau kata benda dalam bahasa Jepang hampir selalu ditulis dengan huruf *kanji* atau *katakana*.

### Verba

Dalam bahasa Jepang, verba dapat dikategorikan ke dalam tiga subkategori berdasarkan bentuk perubahannya. Akan tetapi, subkategori tersebut tidak akan menjadi perhatian dalam makalah ini. Yang terpenting adalah bagaimana cara mengenali verba dalam kalimat. Akar kata verba biasanya ditulis dalam huruf *kanji* kemudian diikuti dengan beberapa huruf *hiragana* yang menunjukkan bentuk perubahannya (bentuk-perubahan-verba selanjutnya akan disebut konjugasi).

### Adjektiva

Adjektiva atau kata sifat dalam bahasa Jepang dibagi ke dalam dua subkategori: adjektiva-i dan adjektiva-na. Adjektiva-i disebut begitu karena dalam bentuk dasarnya adjektiva-i ditulis dengan huruf *kanji* dan diakhiri dengan *hiragana* 'i'. Huruf 'i' tersebut yang akan mengalami perubahan bentuk (perubahan bentuk kata sifat selanjutnya akan disebut deklinasi). Adjektiva-na atau sering juga disebut nomina adjektival mendapat namanya dari kemiripan bentuknya dengan nomina. Adjektiva-na dalam bentuk dasarnya ditulis dengan huruf *kanji* dan akan mendapat akhiran *hiragana* 'na' ketika berfungsi sebagai kata sifat yang langsung menjelaskan kata setelahnya.

### Partikel

Partikel memiliki peran yang penting dalam bahasa Jepang. Di dalam kalimat, partikel menentukan fungsi kata yang mendahuluinya. Apakah suatu kata berfungsi sebagai subjek atau objek ditentukan dari partikel yang mengikuti kata tersebut. Partikel biasanya ditulis dalam

satu atau beberapa huruf *hiragana*.

## C. Satuan Bahasa

Dalam bahasa Jepang ada lima satuan bahasa. Dari yang terbesar, satuan-satuan tersebut adalah sebagai berikut: wacana, paragraf, kalimat, *bunsetsu* (frasa), *tango* (morfem). Dari kelima satuan tersebut, yang menjadi perhatian dalam makalah ini adalah kalimat, *bunsetsu*, dan *tango*.

### Kalimat

Kalimat, atau dalam bahasa Jepang disebut *bun* (文), adalah rangkaian kata-kata yang mengandung makna dan dalam ragam tulis diakhiri dengan tanda titik, tanda tanya atau tanda seru.

### Bunsetsu

*Bunsetsu* (文節) sering kali diterjemahkan sebagai frasa. Akan tetapi, dalam makalah ini penulis akan tetap menggunakan kata *bunsetsu* untuk mengacu pada satuan bahasa Jepang yang berkaitan karena struktur kalimat dalam bahasa Jepang sedikit berbeda.

*Bunsetsu* adalah hasil pemenggalan kalimat dengan aturan tertentu. Aturan yang dimaksud adalah sebagai berikut: 1) satu *bunsetsu* terdiri dari hanya satu morfem bebas dan semua morfem terikat yang mengikutinya, dan 2) morfem bebas tersebut muncul di bagian awal *bunsetsu*.

### Tango

*Tango* (単語) adalah satuan terkecil dalam bahasa Jepang yang tidak dapat diurai lagi tanpa menghilangkan maknanya. Satu *tango* terdiri dari satu morfem, baik bebas maupun terikat.

## III. TEORI MENGENAI POHON SINTAKS

### A. Pohon Sintaks

Pohon sintaks, atau sering pula disebut pohon urai (*parse tree*), adalah pohon terurut dan berakar yang menunjukkan struktur sintaktik suatu string berdasarkan aturan tata bahasa tertentu. Yang dimaksud dengan terurut adalah bahwa daun pohon sintaks menunjukkan urutan sebenarnya unsur-unsur di dalam string. Dalam makalah ini, string yang dimaksud terbatas pada kalimat.

Pohon sintaks biasanya dibangun berdasarkan hubungan konstituen atau hubungan ketergantungan antarunsur di dalam kalimat.

### B. Pohon Sintaks Berbasis Konstituen

Pohon sintaks berbasis konstituen (*constituency-based parse tree*) adalah pohon sintaks yang dibangun berdasarkan hubungan konstituen antarunsur dalam suatu kalimat. Konstituen, dalam istilah linguistik, adalah unsur bahasa yang merupakan bagian dari satuan yang lebih besar. Sebagai contoh, 'anjing saya', 'lebih besar', 'daripada', dan 'anjing lain' adalah konstituen dari 'Anjing saya lebih besar daripada anjing lain.'

Hubungan konstituen diturunkan dari pembagian subjek-predikat sebuah kalimat. Maksudnya adalah bahwa sebuah kalimat dapat dibagi menjadi frasa nominal yang berfungsi sebagai subjek dan frasa verbal yang berfungsi sebagai predikat. Sebagai contoh ‘anjing saya’ adalah frasa nominal dan ‘lebih besar daripada anjing lain’ adalah frasa verbal dari kalimat ‘Anjing saya lebih besar daripada anjing lain.’. Hubungan konstituen juga memiliki bentuk relasi satu-ke-satu-atau-lebih, yang berarti untuk setiap unsur dalam sebuah kalimat, ada satu atau lebih simpul yang berkorespondensi dengan unsur tersebut.

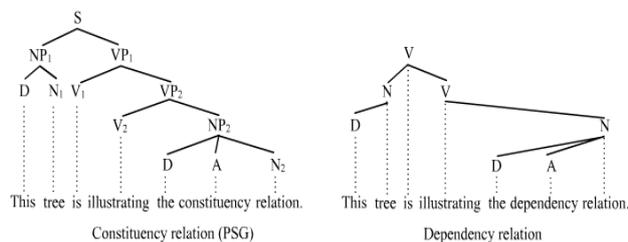
Demikian, pohon sintaks berbasis konstituen memecah suatu kalimat ke dalam subfrasa-subfrasa. Simpul non-terminal dalam pohon sintaks berbasis konstituen adalah jenis frasa atau subfrasa sedangkan simpul terminal atau daun adalah kata di dalam kalimat.

### C. Pohon Sintaks Berbasis Ketergantungan

Berbeda dengan pohon sintaks berbasis konstituen, pohon sintaks berbasis ketergantungan (*dependency-based parse tree*) dibangun berdasarkan hubungan ketergantungan antarunsur dalam suatu kalimat.

Pohon sintaks berbasis ketergantungan menghubungkan kata-kata berdasarkan jenis hubungan mereka. Kata kerja dalam kalimat menjadi akar, kata menjadi simpul dan sisi menunjukkan hubungan ketergantungan. Simpul anak adalah kata yang tergantung pada simpul orangtuanya. Sisi dapat dilabeli berdasarkan jenis hubungannya.

Jika dilihat dari jenis relasinya, hubungan ketergantungan adalah relasi berkorespondensi satu-ke-satu. Artinya, untuk setiap unsur dalam kalimat, hanya ada satu simpul yang berkorespondensi dengan unsur tersebut. Karena itulah, ketika menguraikan kalimat yang sama, pohon sintaks berbasis ketergantungan ini cenderung lebih sederhana daripada pohon sintaks berbasis konstituen.



Gambar 2 Pohon sintaks berbasis konstituen dan pohon sintaks berbasis ketergantungan.

(sumber: <https://upload.wikimedia.org/wikipedia/commons/8/8e/Thistreeisillustratingtherelation%28PSG%29.png>)

## IV. METODOLOGI

### A. Memecah Kalimat Menjadi Bunsetsu

Untuk memecah kalimat bahasa Jepang ke dalam *bunsetsu*-nya, cara berikut dapat dilakukan. Pertama, cari semua morfem bebas yang ada pada kalimat. Perhatikan kalimat berikut.

「彼は優しい人だ。」

*kare wa yasashii hito da.*

Dia adalah orang yang baik.

Morfem bebas dalam kalimat itu adalah 彼, 優しい, dan 人. Pada umumnya, satu atau lebih karakter *kanji* dapat melambangkan sebuah morfem. Dengan pengetahuan tentang kosakata bahasa Jepang yang terbatas, morfem bebas dapat ditentukan dengan cara mengecek kamus bahasa Jepang.

Langkah selanjutnya adalah menentukan *bunsetsu*. Semua karakter setelah morfem bebas hingga karakter sebelum morfem bebas yang lain adalah bagian dari *bunsetsu* morfem bebas tersebut. Perhatikan pemenggalan berikut. Garis miring menunjukkan satuan *bunsetsu*.

「／彼は／優しい／人だ。」

Terlihat bahwa kalimat tersebut memiliki tiga *bunsetsu*. Di bawah ini adalah contoh lain yang lebih rumit.

「日本語能力試験を合格するために、ちゃんと勉強しなければならない。」

*Nihongo nouryoku shiken wo goukaku tame ni, chanto benkyou shinakerebanaranai.*

Untuk lulus JLPT, (aku) harus belajar dengan benar.

Kalimat di atas dipenggal ke dalam *bunsetsu*.

「／日本語能力試験を／合格する／ために、／ちゃんと／勉強しなければならない。」

Kalimat tersebut ternyata memiliki lima *bunsetsu*.

### B. Menentukan Hubungan Ketergantungan

Setelah sebuah kalimat dipecah ke dalam *bunsetsu*-nya, selanjutnya adalah menentukan hubungan antar *bunsetsu* tersebut. Dalam bahasa Jepang hubungan yang dimaksud disebut dengan *kakari-uke* (係り受け). Perhatikan kembali kalimat sebelumnya.

「／彼は／優しい／人だ。」

*kare wa yasashii hito da.*

Dia adalah orang yang baik.

Pada kalimat ini, 彼は (dia adalah) dan 人だ (orang) adalah *kakari-uke* karena orang menjelaskan siapa dia. Kemudian, 優しい(baik) dan 人だ (orang) juga merupakan *kakari-uke* karena karena baik menjelaskan orang yang bagaimana.

Perhatikan contoh lain di bawah ini.

「／私は／山に／登る。」  
*watashi wa yama ni noboru.*  
 Saya mendaki gunung.

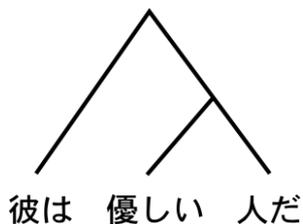
Pada kalimat ini 私は (saya) dan 登る (mendaki) adalah *kakari-uke* karena mendaki menjelaskan apa yang saya lakukan. 山に(gunung) dan 登る(mendaki) juga merupakan *kakari-uke* karena gunung menjelaskan apa yang didaki.

### C. Membuat Pohon Sintaks

Karakter bahasa Jepang yang seperti telah dijelaskan di awal membuat jenis pohon sintaks berbasis ketergantungan lebih cocok untuk digunakan dalam menguraikan kalimat bahasa Jepang. Akan tetapi, karena yang menjadi perhatian dalam makalah ini hanyalah hubungan antar-*bunsetsu*, perlu dilakukan sedikit penyesuaian terhadap pohon sintaks berbasis ketergantungan yang akan digunakan: kata kerja tidak lagi menjadi akar dari pohon, daun dilabeli dengan *bunsetsu* yang ada dalam kalimat, dan hubungan ditunjukkan dengan simpul non-terminal. Sebagai contoh, perhatikan kembali kalimat sebelumnya.

「／彼は／優しい／人だ。」  
*kare wa yasashii hito da.*  
 Dia adalah orang yang baik.

Kalimat di atas jika diubah menjadi pohon adalah sebagai berikut.



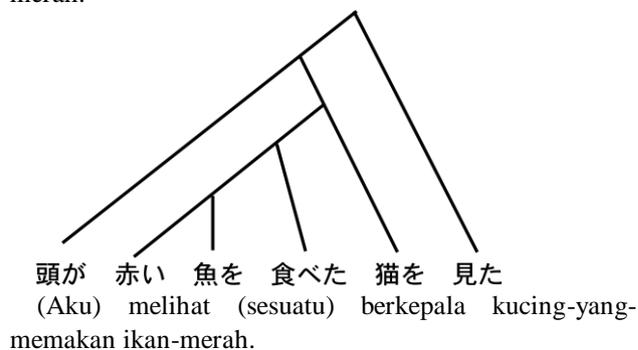
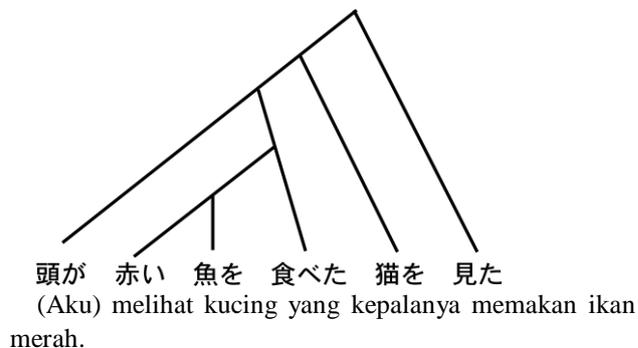
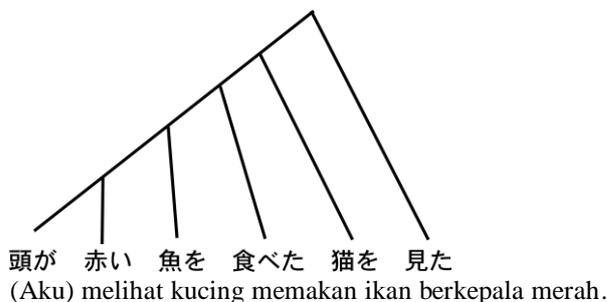
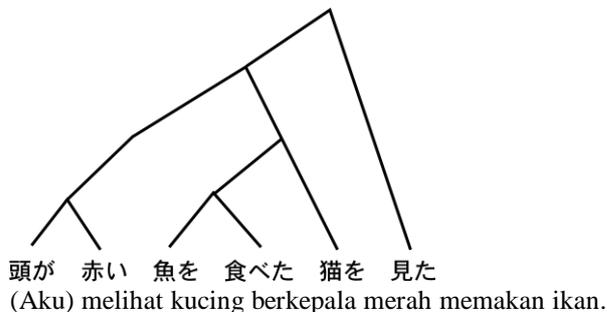
Demikian proses yang dapat dilakukan untuk membentuk pohon sintaks dalam kalimat bahasa Jepang.

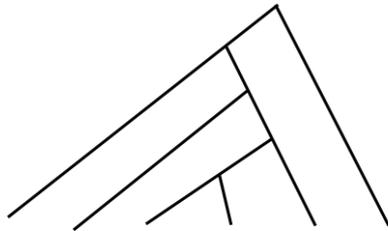
## V. MENGATASI KALIMAT AMBIGU

Perhatikan kalimat di bawah ini.

「頭が赤い魚を食べた猫を見た。」  
*atama ga akai sakana wo tabeta neko wo mita.*

Kalimat di atas adalah kalimat ambigu karena memiliki beberapa pohon sintaks. Bergantung pada pohon sintaksnya, kalimat tersebut juga memiliki beberapa makna semantik yang mungkin. Untuk mengatasi ambigunya kalimat tersebut, maka perlu dicari semua pohon sintaks yang mungkin.





頭が 赤い 魚を 食べた 猫を 見た  
 (Aku) melihat (sesuatu) berkepala kucing-merah-  
 yang-memakan-ikan.

Fleksibilitas bahasa Jepang dalam urutan kata dalam kalimat seringkali menimbulkan kalimat yang ambigu. Untuk mengatasi kalimat yang ambigu tersebut, dapat dibuat pohon sintaks untuk mengetahui hubungan antar-*bunsetsu* dalam kalimat. Dengan mengetahui hubungan itu, maka makna semantik dalam kalimat akan menjadi jelas. Kalimat dengan makna yang sekiranya tidak logis dapat diabaikan. Jika masih terdapat lebih dari satu makna logis, maka makna yang dipilih adalah makna yang paling sesuai dengan konteks kalimat tersebut.

## VI. UCAPAN TERIMA KASIH

Penulis mengucapkan rasa syukur kepada Allah SWT. yang atas rahmat dan kuasanya makalah ini bisa selesai tepat pada waktunya. Penulis mengucapkan terima kasih kepada Bapak Rinaldi Munir sebagai dosen pengampu mata kuliah IF2120 Matematika Diskrit K02 atas bimbingannya. Penulis juga berterima kasih kepada orangtua, keluarga dan teman-teman yang selalu mendukung penulis melalui doa.

## REFERENCES

- [1] Wiyanto, Asul. 2012. *Kitab Bahasa Indonesia*. Yogyakarta:Jogja Bangkit Publisher.
- [2] Mel'cuk, I. 1988. *Dependency syntax: Theory and practice*. Albany: SUNY Press.
- [3] <http://www.guidetojapanese.org/learn/grammar/writing> diakses pada 7 Desember 2016.
- [4] <http://www.genetickanji.com/docs/partsofspeech.asp> diakses pada 8 Desember 2016.
- [5] [http://detail.chiebukuro.yahoo.co.jp/qa/question\\_detail/q1442760009](http://detail.chiebukuro.yahoo.co.jp/qa/question_detail/q1442760009) diakses pada 8 Desember 2016.
- [6] <http://xn--vcs415akpfnn7a.com/> diakses pada 8 Desember 2016
- [7] <http://jisho.org/> diakses pada 7 Desember 2016
- [8] <https://hinoki-project.org/asunaro/output.php> diakses pada 7 Desember 2016
- [9] <https://darkjapanese.wordpress.com/2012/01/04/automatic-sentence-diagramming/> diakses pada 7 Desember 2016
- [10] <http://stackoverflow.com/questions/10401076/difference-between-constituency-parser-and-dependency-parser> diakses pada tanggal 8 Desember.
- [11] <http://www.nltk.org/book/ch08.html> diakses pada tanggal 9 Desember 2016
- [12] <http://japanese.stackexchange.com/questions/11687/how-to-separate-words-in-a-japanese-sentence> diakses pada tanggal 7 Desember 2016

## 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, 8 Desember 2016

ttd

Muhammad Iqbal Al Khowarizmi  
 13515086