

# Rancang Bangun Sistem Pengecekan Ambiguitas Kalimat Berbahasa Indonesia Menggunakan *Harmony Search Algorithm*

Tristi Dwi Rizki  
Fakultas Ilmu Komputer  
Universitas Sriwijaya  
Palembang, Indonesia  
tristidwirizki@gmail.com

Novi Yusliani  
Fakultas Ilmu Komputer  
Universitas Sriwijaya  
Palembang, Indonesia  
novi.yusliani@gmail.com

**Abstrak**—Masyarakat Indonesia sehari-harinya berkomunikasi dan berinteraksi menggunakan bahasa Indonesia. Dalam penggunaannya, masih ada bahasa Indonesia yang tidak sesuai dengan kaidah yang berlaku yaitu masih ada kesalahan-kesalahan dalam penggunaannya. Salah satu jenis kesalahan dalam berbahasa dan berkomunikasi ialah karena adanya ambiguitas. Pada penelitian ini diusulkan sebuah rancangan pengecekan ambiguitas kalimat Berbahasa Indonesia menggunakan Harmony Search Algorithm. Harmony Search Algorithm digunakan untuk menentukan pola pembentuk kalimat. Metode ini menghasilkan sebuah pengecekan kalimat ambigu atau tidak ambigu.

**Keyword**—word sense disambiguation; harmony search algorithm

## I. LATAR BELAKANG

Bahasa Indonesia merupakan bahasa yang dijadikan sebagai bahasa resmi Republik Indonesia dan bahasa persatuan bangsa Indonesia. Masyarakat Indonesia sehari-harinya berkomunikasi dan berinteraksi menggunakan bahasa Indonesia. Meskipun telah terbiasa digunakan, masih ada penggunaan bahasa Indonesia yang tidak sesuai dengan kaidah yang berlaku yaitu masih ada kesalahan-kesalahan dalam penggunaan bahasa baik itu dalam pengucapan maupun penggunaan kata pada suatu konteks tertentu. Salah satu jenis kesalahan dalam berbahasa dan berkomunikasi ialah karena

adanya ambiguitas yang dapat menyebabkan kesalahpahaman dalam memaknai suatu kata atau kalimat.

Ambiguitas merupakan salah satu dari ilmu bahasa Indonesia yang membahas tentang kemungkinan kata atau kalimat yang memiliki dua atau lebih pengertian. *Word Sense Disambiguation*

adalah usaha untuk mengidentifikasi makna kata yang digunakan dalam kalimat tertentu ketika kata memiliki sejumlah makna yang berbeda. Ambiguitas berasal dari bahasa Inggris yaitu *ambiguity* yang berarti suatu konstruksi yang dapat ditafsirkan lebih dari satu arti. Ambiguitas sering juga disebut ketaksaan [1]. Keambiguan dapat mengakibatkan terjadinya lebih dari satu makna pada saat pembicaraan lisan ataupun dalam keadaan tertulis. Saat pembicaraan lisan mungkin dapat diantisipasi dengan pengucapan yang agak perlahan, sedangkan untuk yang tertulis apabila terdapat kekurangan tanda baca maka makna suatu kata atau kalimat menjadi berbeda dari makna yang diinginkan oleh penulis.

Penelitian mengenai pengecekan ambiguitas kalimat sebelumnya telah dilakukan. Penelitian tersebut menggunakan metode *Shift-Reduce Parsing* untuk mengidentifikasi pola kalimat ambigu [2]. Metode ini melakukan identifikasi pola kalimat ambigu dengan mengenali pola kalimat ambigu yang ditentukan oleh kelas kata yang didefinisikan, yaitu kelas kata dari tiap kata dalam kalimat yang digunakan untuk proses *Shift-Reduce Parsing*. Hasil dari penelitian ini menunjukkan bahwa metode *Shift-Reduce Parsing* dapat digunakan secara efektif untuk mengidentifikasi pola kalimat ambigu pada Bahasa Indonesia.

*Prosiding*  
**ANNUAL RESEARCH SEMINAR 2016**  
*6 Desember 2016, Vol 2 No. 1*

ISBN : 979-587-626-0 | UNSRI

<http://ars.ilkom.unsri.ac.id>

Dalam paper ini akan diusulkan sebuah rancangan pengecekan ambiguitas kalimat berbahasa Indonesia menggunakan *Harmony Search Algorithm* yang diharapkan dapat lebih efektif untuk mengidentifikasi sebuah kalimat ambigu atau tidak ambigu. Bagian 2 menjelaskan analisis pra pengolahan kalimat. Bagian 3 menjelaskan analisis kalimat ambigu. Analisis *Harmony Search Algorithm* akan dijelaskan di Bagian 4. Bagian 5 menjelaskan arsitektur perangkat lunak. Dan terakhir Bagian 6 berisi kesimpulan dari seluruh isi paper ini.

## II. PRAPENGOLAHAN KALIMAT

Prapengolahan adalah proses awal mengelola data sebelum pengolahan data dilakukan [3]. Prapengolahan kalimat merupakan tahap awal pemrosesan sebelum dilakukannya proses pengecekan ambiguitas pada kalimat. Adapun tahap prapengolahan kalimat ini terdiri dari *case folding*, *tokenizing* dan *part of speech tagging*. Proses pertama yaitu *case folding*, *case folding* adalah proses mengubah huruf dalam kalimat menjadi huruf kecil. Proses *case folding* hanya dilakukan pada huruf "a" sampai dengan "z", selain huruf itu maka karakter akan dihilangkan.

Proses selanjutnya yaitu *tokenizing*, *tokenizing* adalah proses pemecahan kalimat menjadi kata-kata tunggal yang dipisahkan dengan spasi, tab atau *new line*. Terakhir yaitu proses *part of speech tagging*, *part of speech tagging* adalah proses penandaan kata didalam sebuah kalimat. Pada paper ini digunakan package untuk menandai identitas tiap kata dari kalimat masukan dengan nama NLP\_ITB. Hasil dari proses prapengolahan kalimat dapat dilihat pada gambar dibawah ini.

```
Masukkan Kalimat : Dia berenang di laut mati.  
  
Hasil Proses Case Folding :  
  
dia berenang di laut mati  
  
Hasil Proses Tokenizing :  
  
dia berenang di laut mati  
  
Hasil Proses Part Of Speech Tagging :  
  
dia berenang di laut mati  
PRP VBT IN NNP JJ
```

Gambar 1. Contoh hasil prapengolahan kalimat

Sebelum dilakukan proses *Harmony Search Algorithm* (HSA) hasil dari *part of speech tagging* diolah melalui proses generalisasi terlebih dahulu. Berdasarkan contoh diatas maka hasil generalisasinya dapat dilihat pada gambar 2 dibawah ini.

```
Masukkan Kalimat : Dia berenang di laut mati.  
  
Hasil Proses Case Folding :  
  
dia berenang di laut mati  
  
Hasil Proses Tokenizing :  
  
dia berenang di laut mati  
  
Hasil Proses Part Of Speech Tagging :  
  
dia berenang di laut mati  
PRP VBT IN NNP JJ  
  
Hasil Generalisasi Part Of Speech Tagging :  
  
dia berenang di laut mati  
r v t n n
```

Gambar 2. Contoh hasil setelah proses generalisasi pada *part of speech tagging*

## III. KALIMAT AMBIGU

Kalimat pada umumnya adalah suatu kesatuan bahasa yang berbentuk lisan maupun tulisan yang mengungkapkan sebuah makna atau sebuah pemikiran yang utuh, namun ada kalimat yang mengungkapkan lebih dari satu makna atau pemikiran yang utuh didalamnya. Kalimat yang mengungkapkan lebih dari satu makna tersebut disebut kalimat ambigu. Kalimat ambigu dapat menghambat kelancaran komunikasi karena pihak penerima (pendengar atau pembaca) dapat menafsirkan arti yang berbeda dengan arti yang dimaksud oleh pihak pembicara atau penulis [4]. Contoh kalimat ambigu dapat dilihat pada tabel dibawah ini.

*Prosiding*  
**ANNUAL RESEARCH SEMINAR 2016**  
 6 Desember 2016, Vol 2 No. 1

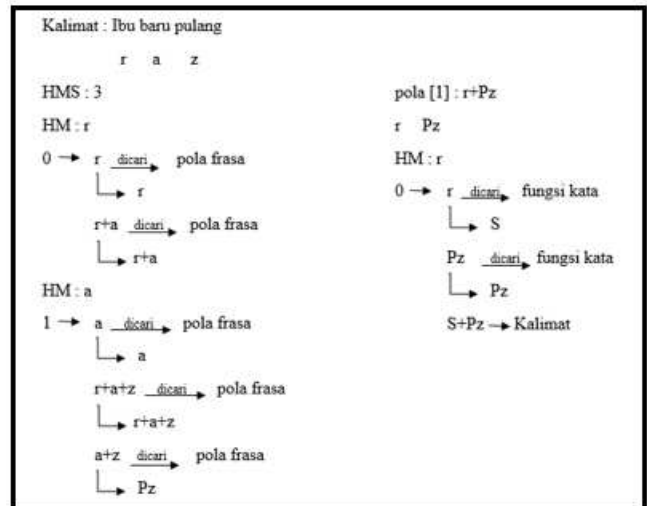
No	Kalimat
1.	Orang malas lewat jalan itu.
2.	Dia berenang di laut mati.

Tabel 1. Contoh kalimat ambigu

Kalimat (1) pada tabel 1 dapat menimbulkan penafsiran lebih dari satu makna (ambigu). Kalimat “Orang malas lewat jalan itu” dapat diartikan dengan hanya orang yang memiliki sifat malas yang akan melewati jalan itu dan orang malas melewati jalan itu karena hal lainnya. Kalimat (2) pada tabel 1 juga dapat menimbulkan lebih dari satu makna (ambigu). Kalimat “Dia berenang di laut mati” dapat diartikan dengan seseorang (Dia) berenang di laut mati yang teletak di Yordania dan seseorang (Dia) berenang di laut kemudian mati.

#### IV. HARMONY SEARCH ALGORITHM

*Harmony Search Algorithm* (HSA) pertama kali diperkenalkan oleh Zong Woo Geem pada tahun 2001 [5]. Ide dasar HSA adalah meniru proses perbaikan harmoni musik yang dilakukan oleh kelompok paduan musik. Ketika kelompok paduan musik melakukan perbaikan pada harmoni musik yang dimainkan, maka akan terdapat tiga kemungkinan pilihan, antara lain memainkan harmoni musik yang terkenal berdasarkan ingatan mereka, memainkan harmoni musik yang serupa dengan harmoni musik yang terkenal namun ada sedikit penyesuaian, atau membuat harmoni musik yang baru [5][6]. Geem (2001) memformulasikan ketiga pilihan ini pada proses optimasi secara kuantitatif. Ketiga komponen tersebut diformulasikan menjadi penggunaan *harmony memory*, penyesuaian nada, dan proses pembangkitan secara *random*. Penggunaan *harmony memory* sangat penting karena *harmony memory* tersebut bisa menjamin bahwa harmoni yang bagus akan dipertimbangkan sebagai elemen-elemen dari vektor solusi yang baru.



Gambar 3. Contoh *Harmony Search Algorithm* pada kalimat Berbahasa Indonesia

Contoh *Harmony Search Algorithm* pada kalimat Berbahasa Indonesia dapat dilihat pada gambar 3 dan berikut ini adalah langkah-langkah dari *Harmony Search Algorithm* (HSA) pada kalimat Berbahasa Indonesia :

1. Inisialisasi masalah dan parameter algoritma;
2. Inisialisasi *harmony memory*;
3. Pengecekan pola frasa dan fungsi kata;
4. Meng-*update harmony memory*;
5. Mengecek kriteria pemberhentian.

Langkah pertama dari *Harmony Search Algorithm* (HSA) adalah inisialisasi masalah dan parameter algoritma. Parameter *Harmony Search Algorithm* (HSA) terdiri dari *Harmony Memory Size* (HMS), *Harmony Memory* dan kriteria pemberhentian.

Langkah kedua dari *Harmony Search Algorithm* (HSA) adalah inisialisasi *harmony memory*. Pada tahap ini dibangkitkan matriks *harmony memory* secara *random* yang berisi populasi hasil generalisasi kalimat masukkan sebanyak HMS.

Langkah ketiga dari *Harmony Search Algorithm* (HSA) adalah pengecekan pola frasa dan fungsi kata kalimat masukkan. Proses pengecekan ini dilakukan dengan memeriksa pola pada masukkan dan database.

Langkah keempat dari *Harmony Search Algorithm* (HSA) adalah meng-*update harmony memory*. *Harmony memory*

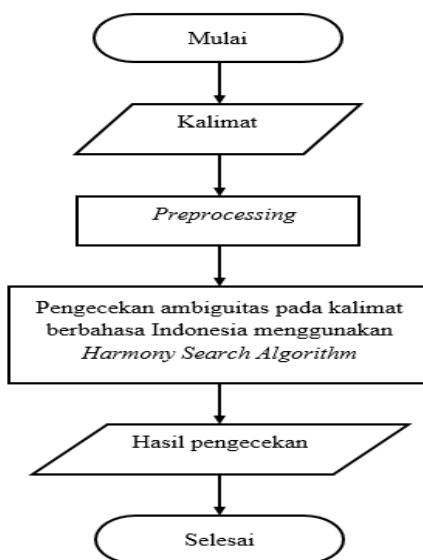
Prosiding  
**ANNUAL RESEARCH SEMINAR 2016**

6 Desember 2016, Vol 2 No. 1

*update* berisi solusi baru hasil pengecekan pola frasa dan fungsi kata pada database.

Langkah kelima dari *Harmony Search Algorithm* (HSA) adalah mengecek kriteria pemberhentian. Apabila kriteria pemberhentian telah tercapai maka iterasi dihentikan, apabila belum tercapai maka kembali ke langkah tiga. Terdapat beberapa macam kriteria pemberhentian antara lain jumlah iterasi maksimal, tidak terjadinya perubahan solusi setelah beberapa iterasi dan lain-lain.

V. ARSITEKTUR PERANGKAT LUNAK



Gambar 4. Alur proses umum perangkat lunak

Gambar 4 diatas menjelaskan proses umum perangkat lunak yang akan dikembangkan pada penelitian tugas akhir ini. Proses dimulai dengan memasukkan kalimat tunggal berbahasa Indonesia, kemudian proses selanjutnya kalimat yang dimasukkan diolah didalam *preprocessing* untuk dilakukan proses *case folding*, *tokenizing*, dan *POST tagging*. Hasil *preprocessing* selanjutnya dilakukan pengecekan ambiguitas menggunakan *Harmony Search Algorithm*, proses pengecekan dilakukan dengan memeriksa pola kalimat, jika kalimat masukkan memiliki lebih dari satu pola kalimat maka

masukkan tersebut merupakan kalimat ambigu dan jika hanya memiliki satu pola kalimat, maka masukkan tersebut bukan kalimat ambigu.

VI. KESIMPULAN

Pengecekan ambiguitas pada kalimat berbahasa Indonesia menggunakan *Harmony Search Algorithm* merupakan pengecekan yang dilakukan pada kalimat tunggal. Kalimat tunggal didapatkan dari penelitian yang terdahulu dan dari pengumpulan yang dilakukan oleh penulis. Diharapkan dengan menggunakan *Harmony Search Algorithm* mampu menghasilkan perangkat lunak yang dapat melakukan pengecekan kalimat ambigu berbahasa Indonesia.

DAFTAR PUSTAKA

- [1] H. Alwi, Soenjono Darhjawidjojo, Hans Lapolowa, dan Anton M, "Tata Bahasa Baku Bahasa Indonesia," Departemen Pendidikan dan Kebudayaan Republik Indonesia, Jakarta, 2003.
- [2] M. Maya Agustin, "Identifikasi Pola Kalimat Ambigu Dalam Bahasa Indonesia Menggunakan *Shift-Reduce Parsing*," Universitas Sriwijaya, Palembang, 2014.
- [3] A. Hatta, Ramadijanti.N, dan Helen, "Ranvang Bangun Sistem Pengelolaan Dokumen-Dokumen Penting Menggunakan Text Mining," Institut Teknologi Sepuluh November, Surabaya, 2010.
- [4] A. Wiyanto, "Kitab Bahasa Indonesia," Jogja Bangkit Publisher, Yogyakarta, 2012.
- [5] Achmad Setiawan dan Ir.Budi Sentosa,M.Sc,P.Hd, "Penerapan Algoritma *Harmony Search* Dalam Penyelesaian *Resource-Constrained Project Scheduling Problem*," Institut Teknologi Sepuluh November, Surabaya.
- [6] Saas Adnan Abed, S.Tiun, dan Nezila Omar, "*Harmony Search Algorithm for Word Sense DisambiguationI*," Knowledge Technology Research Group (KT), Center for Artificial Intelligent (CAIT), Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia, 2015.