

PROCEEDINGS ICCS 2014

International Conference on Computer Systems



APTIKOM



PROGRAM COMMITTEE

Prof. Dr. Ir. Richardus Eko Indrajit M.Sc., MBA., Mphil., MA.

Prof. Dr. Ir. Zainal Arifin Hasibuan, M.Sc., PhD.

Prof. Dr. Salama Manjang, MT.

Drs. H. Achmad Batinggi, MPA.

Drs. Suarga, M.Math., Ph.D.

Dr. Moh. Alifuddin, M.M.

Muhammad Diah Yusuf, Ph.D.

Sofyan S.Thayf, MT.

EDITOR

KETUA PENYUNTING

Dr.Eng. Armin Lawi, S.Si., M.Eng.

WAKIL KETUA PENYUNTING

Dedy Triawan, S.KOM, MMSI

PUBLISHER

Asosiasi Perguruan Tinggi Komputer (APTIKOM) Wilayah IX Sulawesi

Publisher is not responsible for the content

ISSN: 2407-2567

Table of Contents

Nominal Exchange System Design Currency Rupiah's Using Image Pattern Recognition	1
Model Distance Learning Programming Web-Based Algorithm for Improving Student Learning Outcomes High School of Information and Computer Management (STMIK) Handayani	7
Analysis Of Energy Use On The Computer And Design Of Supply Using Solar Power For Remote Schools	13
Designing the Stable Infrastructure for Kernel-based Virtual Machine using VPN-tunneled VNC	18
Andriod-Based Mobile Academic Portal Application	24
Costumer Relationship Management (CRM) To Enhance College Competitiveness	32
A Comparative Study between Two Hybrid Medical Image Compression Methods	38
The Implementation Of Post Editing Method On The Translator Application To Identify The Ambiguous Sentences	45
Enhanced Oriented Bounding Box Method Implementation in Distributed Virtual Environment	51
Operational CRM at Higher Education	55
Design Speed Estimation Of Induction Motor Drives With Dsp (Digital Signal Processing) Based On Neural Network	61
Major Determination Using Artificial Neural Networks For New Prospective Collegian	66
Artificial neural network for measuring the value added products using algorithm backpropagation (case study in cv. Sapo nata de coco durin langkat)	69
The Design Of Virtual Class Android-Based Mobile Video Streaming Using The Real Time Streaming Protocol	74
Analysis of User Experience Quality at IT Telkom's Digital Library Based on Intrinsic Motivation Inventory (IMI) Method	78
Developing the Game Design Document for An Educational Game Named "Bank Sampah"	82
A Proposed Framework for Accounts Receivable Information System (ARIS) Development	88
Smalltalk Application - Learning Programming On Microcontroller For Beginner	94
The Encryption Application Description Image On Mobile Phone Based On Android	99
Design Development Vending Machine With System Rupiah's Currency Returns	103
Implementation Modeling Certainty Factor Determination Of Expert System To Death Victim Not Fair	109

Development of Authoring Tool Flash in the Making of Animation Learning Media	116
File Sharing Data Security Implementation using Access Control List, Screening and Quota	122
The Role of Information and Communication Technology Achieve Trasnparansi and Public Accountability	128
UJARwayfinding: Sign Direction with Augmented Reality Approach	133
Simulation Saccharomyces Cereviciae consumption rate and the Ethanol Formation in Batch Fermentation Process	137
Object Relational Mapping Method in Designing Database Structure for Mosque Management Information System	142
Mobile Web-Based Training (Case Study: Vocational of Education Development Center (VEDC))	146
Maps Dynamic Collaborative Learning In Indonesia Integrated and Open Online Learning (PDITT)	150
Design of Application Framework for Multi-Tenant Asset Management System	154
The Embedded Real Time Video Streaming for Controlling Carrier Robot	161
Haversine and Heron Method Collaboration To Calculate The Area Of A Square Of Irregular Land	164
Finding student learning pattern using cross join Pearson correlation test	167
Database Design for Final Project Monitoring System in Information Management Study Program of Telkom University	173
B-Tree Indexed-File for Genomic Sequence Data Searching	178
Developing Battle Of Etam Earth Game Agent With Finite State Machine (FSM) And Sugeno Fuzzy	184
The Application of C4.5 Method in Determining the Passing of English Proficiency Test (EPT)	188
Application Of Computer Vision For Determination Of Symmetrical Object Position In Three Dimensional Space	192

The Implementation Of Post Editing Method On The Translator Application To Identify The Ambiguous Sentences

Dewi Soyusiawaty

Abstract— In the daily life, there are many words or sentences having ambiguous meaning. The word 'jatuh' could have two different meanings in different sentences. It could mean 1) movement from an upright position to a lower position 2) bankrupt. Ambiguous meaning in the sentences such as 'Pendaftaran anggota baru akan dilaksanakan minggu depan' could mean 1) new member (The new member registration will start next week), and 2) will just start (The member registration will just start next week). The translation system, which is available now, hasn't been able to automatically produce a valid translation output. It needs a collaboration between human and machine. This research discusses the implementation of post editing method to produce a translation as a user means. The research starts by collecting data consisting of basic vocabularies, which have ambiguous meaning, and part of speech both from the source language and the target language. Then, a system need analysis is conducted. This research emphasizes on designing necessary database and user interface related to the interaction between a user and system by post editing process. The implementation is the application produced. The result of the research is the implementation of post editing strategy on the translator application to identify ambiguous sentences to produce valid output sentences as the user means.

Keywords— ambiguous meaning, post editing, translation

1. INTRODUCTION

The Concept of Translation Today, there is no machine-assisted translation system (TMe) that provides fully automatic and valid translation. The one which exists now is translation system involving collaboration between human and machine. The collaborative translation system could be in form of machine-assisted human translation (TMaBMe) or human-assisted machine translation (TMeBMa). In TMaBMe, translation is done by human with computer assistance in form of: (1) the computer dictionary gives the term's meaning based on the context or the topic of the text; (2) the general dictionary computer which is always updated based on the recent terminology; and (3) the computer which could analyze sentences particularly in form of word morphology (often completed with spelling correction).

In the translation, there are two main processes: (1) receive, understand or analyze a source text, and (2) produce, express or synthesize the target text or translation result. Without correctly understanding the source text, what will be

expressed could not be correct. In other words, the translation result doesn't represent the source text or it's not correct. Therefore, the most essential process is correctly understand and analyze the source text. By correctly understanding the source text, the main demand of translation, which is the content or meaning equivalence between the source and the target text, could be achieved. Collaborative translation system which is commonly experimented and got attention of the researchers is TMeBMa system. There are two TMeBMa systems, which are: (1) pre editing type and (2) post editing type. Pre editing type could be described as follows:

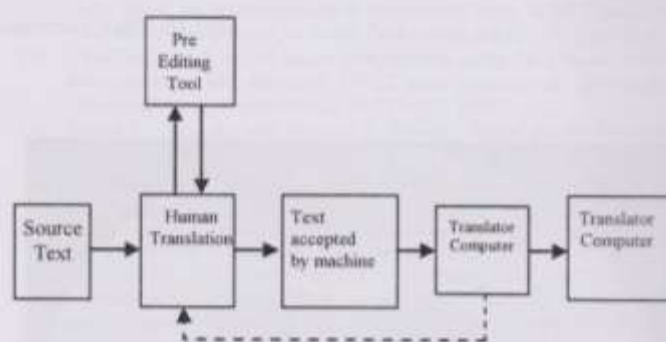


Fig. 1 Pre Editing Type

In this system, a pre editing actor revises the source text to be understandable and acceptable text for the computer as an input, which is then processed to produce the translation result.

In the post editing type, the source text is processed by the translator computer to produce the translation result which has a lot of mistake. This translation result is examined by human translator, and then is edited to be a common translation result. The process could be described as figure 6. In the raw translation result from computer, there are a lot of mistakes which are especially caused by ambiguity in the source text. [1][3][4][5]

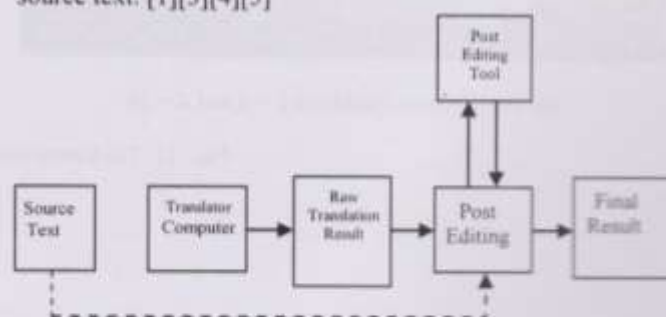


Fig. 2 Post Editing Type

The Ambiguity in the Natural Language

The ambiguity in the natural language could be differentiated, at least, into three kinds, which are (i) category ambiguity, (ii) semantic ambiguity, and (iii) syntactic ambiguity. The category ambiguity is an ambiguity related to parts of speech, as represented by the word 'water' in English language; Look at the example below:

The water is boiling.	[noun]	(1)
They water the plant twice a day.	[verb]	(2)

An example of semantic ambiguity in Indonesian language is the word *genting*. The meaning of *genting* in phrase (3) is different to the meaning in the phrase (4).

Peristiwa genting	(3)
Atap genting	(4)

Another example is the word *kali* which in Indonesian language has three meanings: (1) river; (2) frequency; dan (3) multiply (mathematic operation). The word *kali* in the sentences together with words such as *dalam* (deep), *lebar* (wide), etc, we will know that *kali* means 'river'. Therefore, the language context of the word '*kali*' avoids or solves the ambiguity (disambiguate). Likewise, if the word *kali* comes together with numbers, the word 'equal', etc, we will know that *kali* means 'multiply/ relates to mathematic operation'. From the perspective of providing ambiguity, it needs certain ways to solve the ambiguity/ disambiguate. In case of lexical ambiguity, disambiguating could be done by marking the meaning with a symbol, for example subscript. For instance the word '*kali*' in the sentence "*Bukan tiga kali, tetapi empat kali*", the three meaning mentioned above could be possible. Therefore, it can be given ambiguity solving mark as follows:

- Bukan tiga kali_s, ['sungai']* (river)
- Bukan tiga kali_f, ['frekuensi']* (frequency)
- Bukan tiga kali_m, ['perlakuan matematik']* (multiply/ mathematic operation)

The example of syntactic ambiguity could be found in the Indonesian sentences using the word '*baru*' [8], which could be translated as an adjective describing the noun '*anggota*', also could be translated as an adverb which is related to *minggu depan*. *Pendaftaran anggota baru akan kami lakukan minggu depan* [8]

However, the ambiguity as shown in (3), (4), and also the following example, could be solved by the computer because there are different characteristic that could be revealed. The word *genting* (3) is an adjective while *genting* (4) is a noun. The word *memerah* (9) is a transitive verb while *memerah* (10) is an intransitive verb. The word *mengandung* (11) together with human subject, while *mengandung* (12) relates to non-human subject (thing).

<i>Pekerjaannya memerah susu sapi.</i>	(9)
<i>Wajahnya memerah ketika ia dimarahi.</i>	(10)
<i>Ibu sedang mengandung enam bulan.</i>	(11)
<i>Garam mengandung potassium dan sodium.</i>	(12)

The same case could be found in the potential ambiguous phrase such as : '*wanita dan laki-laki yang tua itu*', which could mean 'those old woman and old man' or ' those woman and an old man'. This phrase is ambiguous because it can be analyzed by both constituent composition and immediate constituent structure:

- ((*wanita dan laki-laki*) yang tua) itu) (those ((woman and man) who are old))
- ((*wanita*) dan (*laki-laki yang tua*) itu) (those ((woman and (old man)))

However, the following ambiguity is still a problem for the computer because it could just analyze a sentence, not context. The ambiguity such as the example (9) – (12) could be solved only using a sentence analysis.

In an interactive system, there is an interaction between computer and human who uses the computer. In this system, (1) a user types a sentence in the source language, (2) the computer asks some necessary questions, (3) the user answers the questions and (4) the computer produces the sentence in the target language without any necessary post editing. The questions and answers solve the ambiguous words or sentence. [2] [6]

Discussion

A. Data

Core Data needed consisting of:

- Basic Vocabularies from a to z. A physical dictionary is a word collection which is arranged and translated to the target language.
- Part of Speech. There are 12 parts of speech, which are nominal, adverbs, verbal, verbs, adjectives, pronouns, numbers, articles, question words, etc. [12]
- Compound words. Words are combined and have one new meaning. As identified, there are more than 300 compound words in Indonesian language and the translation in the target language. For instance *rumah makan*, *kapur sirih*, *air mancur*, *air mata*, etc.
- Homonym. One part of ambiguous varieties is lexical ambiguity, which is a word that is written and spelled the same way but has different meaning. As identified, there are more than 250 homonym words. For instance, the word '*tahu*' could mean understand or a food name made of soy bean. '*Apel*' could mean a fruit or a ceremony. Bisa could mean able to do something or snake's poison, etc.
- Polysemy. A kind of ambiguous word which represents a same meaning of a word but different in the context. As identified, there are more than 150 polysemy words. For example the word '*jatuh*' in the sentence '*Nilai rupiah jatuh*' doesn't mean something which physically falls from an upright position to a lower one, but it means decrease.

TABLE I
Homonym Words

No	Word	Category	Meaning
1	Atlas	Nominal	Map Book
2	Atlas	Nominal	Madras (cloth)
3	Atlas	Nominal	Neck Bone Segment

4	Apel	Nominal	Fruit
5	Apel	Verbal	Ceremony
6	Asal	Verbal	On conditions
7	Asal	Verbal	Careless, random
8			A place (like a basket) made of webbing
9	Bakul	Nominal	Seller
10	Guling	Nominal	Long pillow
11			Going back and forth
12	Guling	Nominal	Baked food
13	Guling	Nominal	Roll over

B. ERD

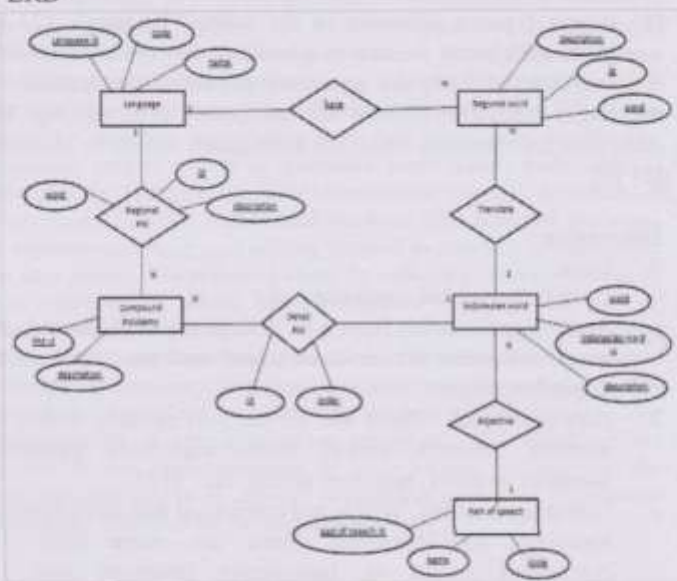


Fig. 3 ERD

Mapping table

1. Language

TABLE II
Table of Language

Couloumn Name	Data Type
Language_id	Int
Code	Char(10)
Name	Varchar(30)

2. Part_of_Speech

TABLE III
Table of Part of Speech

Couloumn Name	Data Type
Part of Speech_id	Int
Code	Char(10)
Name	Varchar(30)

3. Table of Indonesian_Word

TABLE IV
Table of Indonesian Word

Couloumn Name	Data Type
Indonesian_Word_id	Int
Part of speech_id	Char(5)
Word	Varchar(25)
Description	Varchar(255)

4. Table of Regional Word

TABLE V
Table of Regional Word

Couloumn Name	Data Type
Id	Int
Language_id	Int
Indonesian_Word_id	Int
Word	Varchar(25)
Description	Varchar(255)

5. Table of Compound_Polysemy

TABLE VI
Table of Coumpound Polysemy

Couloumn Name	Data Type
PM_id	Int
Description	Varchar(255)

6. Table of Detail_Compound_Polysemy

TABLE VII
Table of Detail Compound Polysemy

Couloumn Name	Data Type
Id	Int
PM_id	Int
Indonesian_Word_id	Int
Order	Int

7. Tabel of Regional PM

TABLE VIII
Table of Regional PM

Couloumn Name	Data Type
Id	Int
Language_id	Int
PM_id	Int
Word	Varchar(25)
Description	Varchar(255)

II. IMPLEMENTATION

1. Data Inputing Page

Data Inputing Page is prepared to input data of Indonesian words both homonym and not homonym with the regional language translation, polysemy words or compound words of Indonesian language, also, with their translation.

Fig. 4 Data Inputting of Indonesian Words

The following figure shows the setting of ambiguous words data or homonym. For example, the word 'Apel' could mean a noun (N) or a verb (V) with description of 'buah' (fruit) and 'upacara' (ceremony).

Fig. 5 The Data setting of Ambiguous Words

Figure 6 shows the regional words data from Indonesian words.

Fig. 6 The setting of Regional words

To input polysemy data or compound words is shown as following figure. A compound word could be inputted if only each word combined in that compound word has already been inputted on the inputting page in the previous Indonesian words data.

Fig. 7 The Data setting of Indonesian Polysemy/Compound

After the data of compound words is inputted on the compound word page, then the translation of regional language words could be inputted. The following figure shows the page where regional language compound words are inputted.

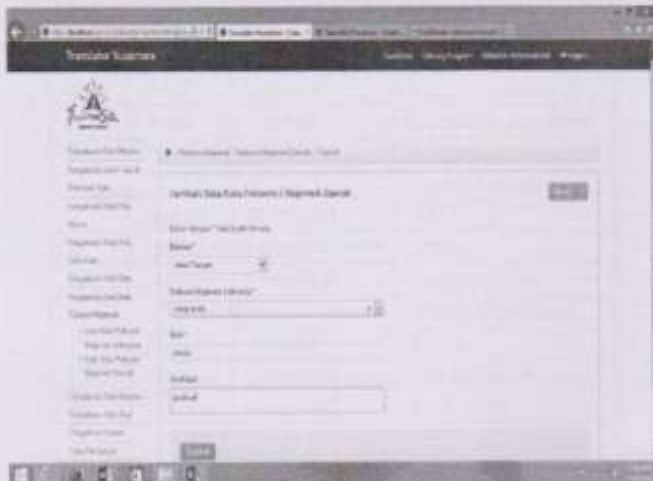


Fig. 8 The Data setting of Regional Compound Polysemy

The page below shows the list of compound words of the regional language.



Fig. 9 The Setting of Regional Compound Words

2. Translation Page

The implementation on the translation page is as follows:



Fig. 10 Translation Page of Homonym Word

The word 'tahu' is one of homonym words. The implementation of post editing method is system gives a

different sign on the translation page showing that word could have more than one alternative translations or ambiguous. For example the word 'tahu' as a verb and a noun.



Fig. 11 Alternative translation or Ambiguous meaning

A user could select a translation option based on what he/she actually means. After the editing process, the system will apply the option in the translation result. The same thing happens in ambiguity example in the compound words, as shown on the following figure :



Fig. 12 Translation of Compound Words

The word 'panjer' is the first alternative shown by the system as the word 'uang muka' inputted. On alternative translation page or ambiguous meaning, there are two optional translations which are 1) panjer and 2) 'dhuwit' dan 'muka'. On the second option, the system shows translation of each word if they aren't actually a compound word but independent words.



Fig. 13 Alternative Result

A user could determine the meaning as he/she means. The system will show the translation as selected.

III CONCLUSION

- a. Post editing method is one of solutions on developing translator application from one language to another to identify ambiguous sentences.
- b. The implementation of post editing method on the translator application aims to get translation result as a user means.

REFERENCES

- [1] Andreas, Steve and Faulkner, Charles, "NLP The New Technology of Achievement", Nicholas Brealey Publishing, 1999.
- [2] Chaer, Abdul, "Pengantar Semantik Bahasa Indonesia", Jakarta: Rineka Cipta, 2009.
- [3] Hurford, James R. and Brendan Heasley, "Semantics a Coursebook", Cambridge University Press, 1987.
- [4] Lyons, John, "Linguistics Meaning: an introduction", Cambridge University Press, 1995.
- [5] Patterson, Don W, "Introduction to Artificial Intelligence and Expert Sistem", Prentice Hall Internasional, Inc, 1990.
- [6] Ramlan, "Penggolongan Kata", Yogyakarta : Andi Offset, 1985.