

AN ENGLISH-TO-INDONESIAN
MACHINE TRANSLATION
SYSTEM

By

MOHAMMAD TEDUH ULINIANSYAH

Bachelor of Science

Institute Technology of Shibaura

Tokyo, Japan

1991

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
July, 1998

OKLAHOMA STATE UNIVERSITY

AN ENGLISH-TO-INDONESIAN
MACHINE TRANSLATION
SYSTEM

Thesis Approved:

W. Vick Stuart

Thesis Adviser

John Hateliff

D. E. Hedrick

Wayne B. Powell

Dean of the Graduate College

ACKNOWLEDGEMENTS

I would like to express my gratitude to my major adviser, Dr. W. Nick Street for his continuous guidance and help, and to other committee members, Dr. G.E. Hedrick and Dr. John Hatcliff for their help and good will while I was writing this thesis.

I would like to acknowledge the financial support I have received from the Indonesian Government during my graduate study. This financial support is administered by the Minister for Research and Technology/the Agency for the Assessment and Application of Technology in the framework of the implementation of the Science and Technology for Industrial Development Program.

Finally, I wish to express my deepest love and affection for my wife and daughter who never failed in their support and love for me during my graduate studies in the Department of Computer Science.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
Motivation	1
Method and Limitations	3
Organization of Study	3
II. LITERATURE REVIEW	5
History	5
Types of MT	6
General Problems in MT	11
Ambiguity	11
Lexical and Structural Mismatches	12
Multiword Units: Idioms and Collocations	12
LINK Grammar System	12
III. BRIEF DESCRIPTION OF THE INDONESIAN LANGUAGE	20
IV. SYSTEM	21
Overview	21
Operation	22
System Performance	27
V. EVALUATIONS, CONCLUSIONS, AND SUGGESTED FUTURE RESEARCH	34
Evaluation	34
Conclusions	35
Suggested Future Research	35
SELECTED BIBLIOGRAPHY	37
GENERAL REFERENCES	39
APPENDICES	40

Chapter	Page
APPENDIX A - USER'S MANUAL	41
APPENDIX B - DESCRIPTIONS OF CONNECTORS	44
APPENDIX C - SOURCE AND TRANSLATED TEXTS	47

LIST OF TABLES

Table	Page
I. A Simple Dictionary for LINK Grammar	15
II. Optional Expression in the Dictionary	16
III. Performance of the Analysis System	28
IV. Performance of the Generation System (Calculations Based on the Results of the Analysis System)	31
V. The Overall Performance of the System	31

LIST OF FIGURES

Figure	Page
1. Direct Approach	7
2. Interlingua Approach	8
3. Transfer Approach	9
4. Linking Requirement of Words a, the, cat, snake, Mary, ran, and chased	13
5. A Satisfied Linking Requirement	14
6. An Unsatisfied Linking Requirement	14
7. Simplified Form of Linkages	15
8. Links for “The big snake the black cat chased bit Mary”	16
9. Overview of the Translation System	22
10. An Example of the Analysis Result	23
11. A Sample Input File	42

CHAPTER I

INTRODUCTION

Motivation

The term machine translation (MT) is the traditional and standard name for computerized systems responsible for the production of translations from one natural language into another, with or without human assistance. MT has been a very popular research area in the field of natural language processing (NLP). NLP is a field in artificial intelligence that focuses on aspects of human languages. The aim of MT is to automate all or part of the process of translating from one human language to another. It was one of the earliest applications suggested for digital computers. However, turning the dream into reality has turned out to be much harder, and in many ways, a much more interesting task than it first appeared.

We can see the importance of MT from two points of view: *social or political* importance and *commercial* importance [Arnold, 1994]. The social or political importance of MT arises from the *socio-political* importance of translation in communities where more than one language is generally spoken. The community could use a single common 'lingua franca' rather than translation. However, this is not an attractive alternative because it involves the dominance of the chosen language. This is a disadvantage to speakers of the other languages. This would also raise the prospect of the other languages becoming second-class and ultimately disappearing. Since the loss of a language often involves the disappearance of a distinctive culture and a way of thinking, this is a loss that should matter to everyone. Therefore, translation is necessary for

communication - for ordinary human interaction and for gathering the information one needs to play a full part in society. In other words, it seems as though automation of translation is a social and political necessity for modern societies which do not wish to impose a common language on their members. The commercial importance of MT is a result of related factors. First, translation itself is commercially important. An English speaker would buy a book written in English rather than one which is written in Japanese. Secondly, translation is expensive. Translation is a highly skilled job requiring much more than mere knowledge of a number of languages. In some countries, translators' salaries are comparable to that of other highly trained professionals.

Scientifically, MT is interesting because it is an obvious application and testing ground for many ideas in computer science, artificial intelligence, and linguistics. Some of the most important developments in these fields have begun in MT.

Philosophically, MT is interesting because it represents an attempt to automate an activity that can require the full range of human knowledge; that is, for any piece of human knowledge, it is possible to think of a context where the knowledge is required.

Surprisingly, universities in Indonesia show little interest in doing research on natural language processing. The Agency for the Assessment and Application of Technology, a government agency, pioneered the research when it joined an international project called Multilingual Machine Translation System (MMTS). The project began in 1987, involving Indonesia, Japan, Malaysia, Thailand and China. The project's goal was to build a machine translation system which can translate sentences among the participants' languages. As the project's name indicates, the system uses the interlingua method (see chapter II) to translate sentences. The project was ended in 1995. The same

agency is now conducting a joint research project in machine translation which is funded by the United Nations.

Method and Limitations

The main objective of this study is to design and build a uni-directional translation software which is able to translate English sentences into the Indonesian language. Since this project is of a large size, this paper only concentrates on the beginnings of the overall system with further development left to future research. The system is developed based on the so-called 'direct' approach (see chapter II). It uses LINK grammar system version 2.0 (see chapter II) for analyzing the input sentences. In order to keep the design of this system at a manageable level, its scope is limited to selected domains of input texts. The selected domains of input texts are politics, economics, and popular technology. Two articles from the *Asiaweek* magazine and one article from the *Inside Indonesia* journal were used to test the capability of the system. The two articles taken from the *Asiaweek* magazine were about economy and popular technology while the article taken from the *Inside Indonesia* journal was about politics.

Organization of Study

Chapter I deals with the motivation for the system and the method of construction. Chapter II includes a literature review discussing papers in the following areas: history, types of MT, general problems faced in MT, and the LINK grammar system. Chapter III discusses the Indonesian language briefly. The system operation and performance on the

computer is shown in Chapter IV. Chapter V presents the evaluation of the system, conclusions of the study, and suggestions for future research.

CHAPTER II

LITERATURE REVIEW

History

The very beginning of research on machine translation can be traced to an idea proposed by Warren Weaver of the Rockefeller Foundation in 1949. He suggested that if the code-deciphering techniques developed during the second World War were used, computers would be able to recognize the fundamental aspects of all known languages [Weaver, 1949]. The analogy of translation and decoding may sound like a simplification of the difficult problems faced in MT. However, it had the virtue of turning an apparently difficult task into one that could be approached with the emergent computer technology. The suggestion sparked a significant amount of interest and research, and by the early 1950s, there was a large number of research groups working in Europe and the USA representing a significant financial investment. But, despite some success, there was widespread disappointment on the part of funding authorities at the return on their investment and doubts about the possibility of automating translation in general. The doubts reached a peak when in a report, the Automatic Language Processing Advisory Committee (ALPAC) concluded that there was no shortage of human translators and that there was no immediate prospect of MT producing useful translation of general scientific texts [ALPAC, 1966]. This report led to the virtual end of government funding in the USA. Worse, it led to a general loss of morale in the field, as early hopes were perceived to be groundless. It was not until the late 1970s that MT research underwent something of a renaissance.

The Commission of the European Communities (CEC) purchased the English-French version of the Systran (System translation) system [Toma, 1976; 1977] in 1978, a greatly improved descendent of the earliest systems developed at Georgetown University. The CEC also commissioned the development of a French-English version and an Italian-English version. At about the same time, there was a rapid expansion of MT activity in Japan, and the CEC also began to set up what was to become the Eurotra project, building on the work of the GETA (Groupe d'Etudes pur la Traduction Automatique) [Vauquois, 1975] and SUSY (Saarbrücker Übersetzungssystem 'Saarbrücken Translation System') [Maas, 1977; 1987] groups. This was perhaps the largest and certainly among the most ambitious research and development projects in natural language processing. The aim was to produce a 'pre-industrial' MT system of advanced design (what we call a linguistic knowledge system) for the European communities languages. Also in the late 1970s, the Pan American Health Organization (PAHO) began development of a Spanish-English MT system (SPANAM) [Vasconcellos, 1988]. Similarly, the United States Air Force funded work on the METAL (Mechanical Translation and Analysis of Language) [Bennet and Slocum, 1985; Slocum, 1987] system at the Linguistics Research Center at the University of Texas in Austin. In addition, the results of work at the TAUM (Traduction Automatique de l'Université de Montréal) group led to the installation of the Météo system [Chandioux, 1976].

Types of MT

The process of machine translation can be divided into three distinct stages:

- (i) The analysis of source language text.
- (ii) The transfer of the expression of the analysis' results into an expression related to the target language.
- (iii) The generation of the text in the target language [Nagao, 1989, p. 6].

Broadly speaking, there are three types of MT:

- (i) Direct translation approach where the MT system is designed in all details specifically for one particular pair of languages in one direction.
- (ii) Interlingua approach, which assumes the possibility of converting texts to and from a 'meaning' representation common to more than one language.
- (iii) Transfer approach, where the MT system involves three stages. The first stage converts texts into intermediate representations; the second stage converts the representations into equivalent representations of the target language; and the third step generates the target language [Hutchins & Somers, 1992, p. 4].

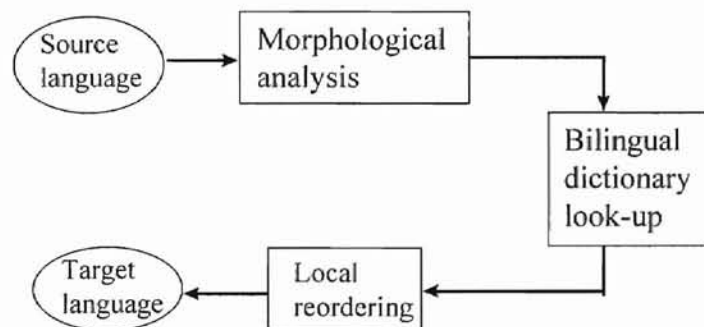


Figure 1. Direct Approach [Hutchins & Somers, 1992, p. 72]

Direct translation is the traditional approach in MT systems. This approach is designed for an MT system which handles one particular pair of languages in all details. The approach assumes that the vocabulary and syntax of source language texts need not

be analyzed any more than the resolution of ambiguities, the correct identification of appropriate target language expressions and the specification of target language word order. Therefore, if the sequence of words in the source language is sufficiently close to an acceptable sequence of target language words, the system does not need to identify the syntactic structure of the source language text [Hutchins, 1986]. The PAHO's (Pan American Health Organization) SPANAM (Spanish-English MT system) is an example of an MT system which employs the approach [Vasconcellos, 1988]. However, some direct MT systems also employ full sentence analysis such as the MT system developed in Wayne State University [Josselson, 1972]. Furthermore, there is a greater degree of modularity in Systran (System translation) [Toma, 1976;1977]. Now, the difference between MT systems which use the direct approach and the transfer approach is narrowing.

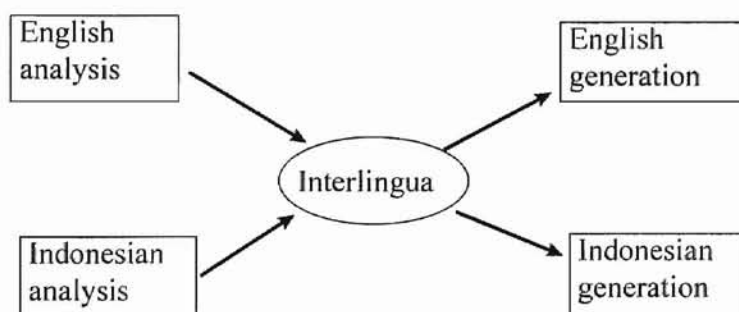


Figure 2. Interlingua Approach [Hutchins & Somers, 1992, p. 74]

The interlingua approach assumes that it is possible to convert source language texts into semantico-syntactic representations common to more than one language. In interlingual systems, the process of translating source language texts to target language texts is carried out in two distinct and independent stages. In the first stage, the analysis

system analyzes the source language text and produces an interlingual representations. In the second stage, interlingual forms serve as the sources for the generation system to generate target language texts. Procedures for source language analysis are intended to be source language-specific and not devised for any particular target language in the system. Similarly, target language generation is intended to be target language-specific.

Some examples of MT systems using the interlingua method are DLT (Distributed Language Translation) [Witkam, 1983] and Rosetta [Appelo and Lansbergen, 1986].

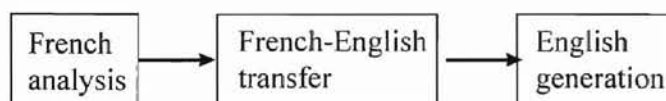


Figure 3. Transfer Approach [Hutchins & Somers, 1992, p. 75]

In the transfer approach, the process of translating source language texts into target language texts is done in three stages. In the first stage, the system converts source language texts into source language transfer representations. The transfer representation is an abstract representation of a natural language text. In the second stage, it converts the source language transfer representations into target language transfer representations. Finally, in the third stage, the system produces the final target language texts from these representations. Whereas the interlingual approach necessarily requires complete resolution of all ambiguities and anomalies of source language texts so that translation should be possible into any other language, in the transfer approach only those ambiguities inherent in the language in question are handled. Differences between languages are handled during the transfer process, that is, problems of mismatch between

source language and target language lexical ranges are resolved in the transfer component. Generally, there are two processes involved during the transfer process. The first process is lexical transfer where the source language words are replaced by their target language meanings. The second transfer is structural transfer where the system changes the form of the representations according to the grammar of the target language by using transfer rules.

The transfer method is employed in GETA (Groupe d'Etudes pur la Traduction Automatique) [Vauquois, 1975], SUSY (Saarbrücker Übersetzungssystem 'Saarbrücken Translation System') [Maas, 1977; 1987], TAUM's (Traduction Automatique de l'Université de Montréal) Météo [Chandioux, 1976], and METAL (Mechanical Translation and Analysis of Language) [Bennet and Slocum, 1985; Slocum, 1987].

The basic difference between the two indirect approaches (transfer and interlingua approach) and the direct approach lies in the configuration of dictionary and grammar data. In the direct systems, the major component is a single source language-target language bilingual dictionary incorporating not only information on lexical equivalents but also all data necessary for morphological and syntactic analysis, transfer and synthesis. In indirect systems, this information is dispersed among separate source language and target language dictionaries, separate source language and target language grammars, and either the interlingua vocabulary and syntax, or the source language-target language transfer dictionary (of lexical equivalences) and a 'grammar' of source language-target language structure transfer rules [Hutchins, 1986].

General Problems in MT

The reasons why MT is difficult to develop can be categorized under three headings:

- Problems of ambiguity.
- Problems that arise from structural and lexical differences between languages.
- Multiword units like idioms and collocations.

Ambiguity

It is common that a word in most languages has more than one meaning. There are two types of ambiguity concerned here. When a word has more than one meaning, it is said to be lexically ambiguous. When a phrase or sentence can have more than one structure, it is said to be structurally ambiguous. Ambiguity is a pervasive phenomenon in human languages. It is very hard to find words that are not at least ambiguous in the two ways above. The ambiguity problems faced in MT is hard because ambiguities 'multiply'. A sentence with two ambiguous words may be ambiguous in four ways. A sentence with three ambiguous words may be ambiguous in eight ways and so on. In this way, one can get very large numbers indeed. For example, a sentence consisting of ten words where each of its words is ambiguous in two ways, and with just two possible structural analyses, could have $2^9 \times 2^2 = 2^{11} = 2048$ different analyses. The number of analyses can be problematic since one may have to consider all of them, rejecting all but one [Arnold, 1994].

Lexical and Structural Mismatches

This problem arises when two different human languages use different perspectives to classify the world. For example, English chooses different verbs for the action/event of putting on and the action/state of wearing. Japanese does not make this distinction, but differentiates according to the object that is worn [Arnold, 1994].

Multiword units: Idioms and Collocations

Idioms are expressions whose meanings cannot be completely understood from the meanings of the component parts. In many cases, a natural translation for an idiom will be a single word. One problem with sentences which contain idioms is that they are typically ambiguous in the sense that either a literal or idiomatic interpretation is generally possible. This problem does not arise with all idioms. Some are completely frozen forms whose parts always appear in the same form and in the same order. Examples are phrases like *in fact*, or *in view of*. However, such idioms are by far the exception. A typical way in which idioms can vary is in the form of the verb which changes according to tense, as well as person and number [Arnold, 1994]. For example, with *pull somebody's leg* ('to play a joke on somebody'), one gets *He is pulling/pulls/pulled somebody's leg*, and *They are pulling/pull/pulled somebody's leg*.

LINK Grammar System

A group of researchers at Carnegie Mellon University have been developing a new technique to parse natural language processing. The system is called LINK grammar

system [Sleator, 1991]. In his observation, Melcuk (1988) claims that most sentences of most natural languages have the property that if arcs are drawn connecting each pair of words that relate to each other, then the arcs will not cross. The working of LINK grammar is based on this observation. The well-known phenomenon, called planarity, is the basis of link grammars as a new formal language system [Sleator, 1991].

A link grammar consists of a set of words (the terminal symbols of the grammar), each of which has a linking requirement. A sequence of words is a sentence of the language defined by the grammar if there exists a way to draw arcs among the words so as to satisfy the following conditions:

- Planarity: The links do not cross (when drawn above the words).
- Connectivity: The links suffice to connect all the words of the sequence together.
- Satisfaction: The links satisfy the linking requirements of each word in the sequence [Sleator, 1991, p. 1].

The linking requirements of each word are contained in a dictionary. Figure 4 which shows a simple dictionary for the words *a*, *the*, *cat*, *snake*, *Mary*, *ran*, and *chased*, illustrates the linking requirements. The linking requirement of each word is represented by the diagram above the word.

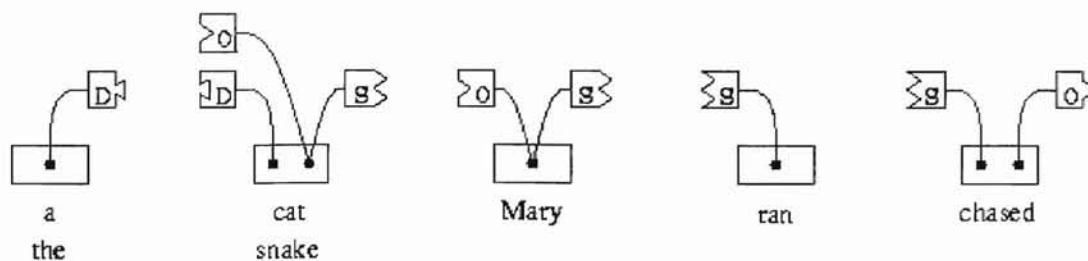


Figure 4. Linking Requirements of Words *a*, *the*, *cat*, *snake*, *Mary*, *ran*, and *chased* [Sleator, 1991, p. 1].

Each of the intricately-shaped, labeled boxes is a connector. A connector is satisfied by "plugging it into" a compatible connector (as indicated by its shape). If the mating end of a connector is drawn facing to the right, then its mate must be to its right facing to the left. Exactly one of the connectors attached to a given black dot must be satisfied (the others, if any, must not be used). Thus, *cat* requires a D connector to its left and either an O connector to its left or a S connector to its right. Plugging a pair of connectors together corresponds to drawing a link between that pair of words. Figure 5 shows how the linking requirements are satisfied in the sentence "*The cat chased a snake*".

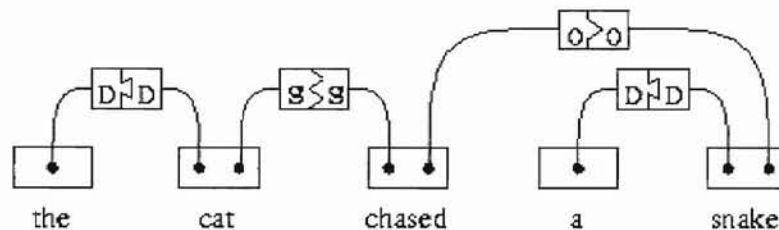


Figure 5. A Satisfied Linking Requirement [Sleator, 1991, p. 1].

The sentences "*Mary chased the cat*", and "*The cat ran*" are also sentences of this grammar. The sequence of words: "*The Mary chased cat*" is not in this grammar. Any attempt to satisfy the linking requirements leads to a violation of one of the three rules developed by Sleator (1991).

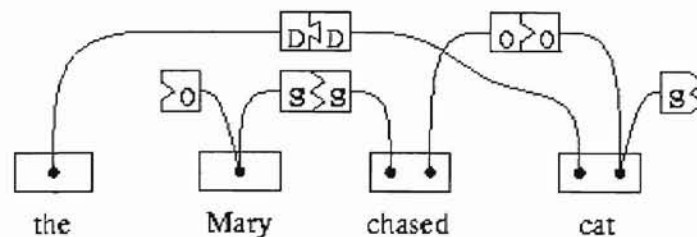


Figure 6. An Unsatisfied Linking Requirement [Sleator, 1991, p. 2].

Similarly, sentences "*ran Mary*" and "*cat ran chased*" are not part of this language.

A set of links that prove that a sequence of words is in the language of a link grammar is called a linkage. The following picture shows a simplified form of the diagram showing that "*the cat chased a snake*" is part of this language:

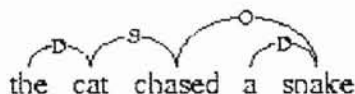


Figure 7. Simplified Form of Linkages [Sleator, 1991, p. 2].

There is a dictionary file which consists of linking requirements expressions. The dictionary in Table I encodes the linking requirements of the previous example.

Table I.

A Simple Dictionary for LINK Grammar [Sleator, 1991, p. 2].

Words	Formula
a the	D+
snake cat	D- & (O- or S+)
Mary	O- or S+
ran	S-
chased	S- & O+

The linking requirement for each word is expressed as a formula involving the operators *&*, *or*, *parentheses*, and connector names. The + or - suffix on a connector name indicates the direction (relative to the word being defined) in which the matching connector (if any) must lie. The *&* of two formulas is satisfied by satisfying both the formulas. The *or* of two formulas requires that exactly one of its formulas be satisfied.

The order of the arguments of an & operator is significant. The farther left a connector is in the expression, the nearer the word to which it connects must be. Thus, when using *cat* as an object, its determiner (to which it is connected with its D- connector) must be closer than the verb (to which it is connected with its O- connector).

The grammar also provides ways to express optionality by using curly brackets {}, and the symbol "@" is used to allow one or more connectors of a type to be attached.

Table II shows the dictionary which illustrates the use of {} and the symbol @.

Table II.

Optional Expressions in the Dictionary [Sleator, 1991, p. 3].

Words	Formula
a the	D+
snake cat	{@A-} & D- & {B+} & (O- or S+)
chased bit	S- & (O+ or B-)
ran	S-
big green black	A+
Mary	O- or S+

The sentence "*the big snake the black cat chased bit Mary*" is in the language defined by this grammar because of the following links:

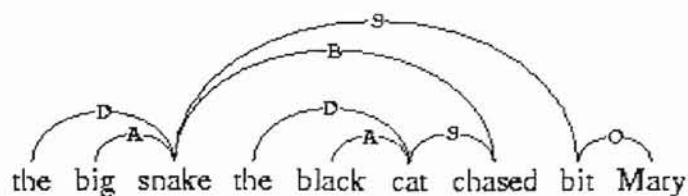


Figure 8. Links for "The big snake the black cat chased bit Mary"
[Sleator, 1991, p. 3].

In this case, there is a unique linkage satisfying all the requirements. The connector A links adjectives to following nouns and connector B connects a transitive verb back to its object in relative clauses.

The LINK grammar system has no explicit notion of grammatical constituents (a word, phrase, or clause forming part of a larger construction). In LINK grammar, constituents can be seen to emerge as contiguous connected collections of words attached to the rest of the sentence by a particular type of link. For example, in Table II above, *S* links always attach a noun phrase (the connected collection of words at the left end of the link) to a verb (on the right end of the link). *O* links work in a similar fashion. In these cases, the links of a sentence can be viewed as an alternative way of specifying the constituent structure of the sentence. The LINK grammar system can perhaps be viewed as a technique for parsing using the assumption that there exists an underlying dependency structure.

The analysis system of this project is based on the LINK grammar system version 2.0. LINK grammar captures many phenomena of English grammar. It handles noun-verb agreement, questions, imperatives, complex and irregular verbs (i.e. *wanted*, *go*, *denied*, etc.), different types of nouns (mass nouns, those that take *to*-phrases, etc.), past- or present-participles in noun phrases, commas, a variety of adjective types, prepositions, adverbs, relative clauses, possessives, and many other things. After some modification, the dictionary for analysis now contains 58,910 words. Individual forms of verbs and nouns such as "chase", "chases", "chased", and "chasing" are counted as separate words. The number of "stem" words is probably about 30,000. The LINK grammar system is

available via anonymous ftp: /afs/cs/user/sleator/public/link-grammar on host ftp.cs.cmu.edu.

The LINK grammar has the following advantages:

- The grammar is lexical; that is, there is a given definition for each word of the lexicon describing how it can be used in a sentence. This makes construction of a large grammar easier and the grammar can easily be constructed incrementally.
- Words that are associated semantically and syntactically are directly linked. This makes it easy to enforce agreement and to gather statistical information about the relationship between words.
- In English, whether or not a noun needs a determiner is independent of whether it is used as a subject, an object or even if it is part of a prepositional phrase. The LINK grammar system takes advantage of this orthogonality. For example, the word “cat” can be used in 369 different ways in the analysis dictionary. A compact LINK grammar formula captures a large number of possibilities and can easily be written and comprehended [Sleator, 1991, p. 3].

The reasons to use the LINK grammar system as a tool to parse the English language are as follows:

- It is not only able to parse grammatical sentences, but it also able to parse ungrammatical sentences such as conversations by using a *null-link* to allow connections between any pair of adjacent words regardless of their dictionary definitions [Grinberg, 1995].
- It is easy to create grammar rules to parse input sentences.

- It is available for free for a research purpose and the system can be obtained via anonymous ftp through the Internet.
- The system comes with good documentation.
- The system is written in C language, making it easier to modify the code.

CHAPTER III

BRIEF DESCRIPTION OF THE INDONESIAN LANGUAGE

The language of Indonesia is called Bahasa Indonesia. It is based on Malay (Bahasa Melayu), one of the more than two hundred local languages spoken in the East Indies [MacDonald, 1967]. Syntactically, Indonesian sentences and English sentences share similar surface structure. The order of words in an Indonesian sentence is like that of an English sentence [Kwee, 1993]. An example is given below.

English sentence: *The professor/reads/a book.*

Indonesian sentence: *Profesor itu/membaca/sebuah buku.*

As we see in the example above, the word order is similar except for the place of the word “*itu*”. The word *itu* in the Indonesian language has several functions. It can be used as an adjective, a pronoun, or a definite article such as in the example above.

The Indonesian language does not have tenses to express time. Instead, it uses ‘functional words’ such as *kemarin* (‘yesterday’), *besok* (‘tomorrow’), etc. to express the time when events/actions occurred. Unlike English and other languages, verbs in Bahasa Indonesia do not take on different forms showing number, person, tense, or mood. Also, the Indonesian language does not have gender pronouns as in English. The most difficult part of the Indonesian language is that it has affixes to change a word from a noun to a verb or adjective, and vice versa. The combinations of infixes, suffixes, and prefixes in the Indonesian language are so many that it is almost impossible to create a certain formula to handle them.

CHAPTER IV

SYSTEM

The system is coded in the C language and consists of a total of 25 modules. The user's manual may be found in the appendices, but since the system is of a large size, the source codes are not presented in this paper. The actual programs are on diskette and may be obtained from the Computer Science Department of Oklahoma State University. The computer on which the system runs is a Sun Sparc 5.

Overview

Although there are many techniques for creating machine translation, this study presents research on creating machine translation with a direct system, as described in Chapter II. The direct system has several advantages. First, it is designed to handle one-way translation, covering all aspects of the source language and the target language. Secondly, it can handle complex forms of input sentences. Moreover, as we have seen, the surface structure of the English language is similar to that of the Indonesian language. Therefore, the direct approach was chosen as a basis to develop the system. The disadvantage of the direct system is that the whole system must be re-written if one wants to add another human language for the system. The overall system of the translation software is illustrated in Figure 9.

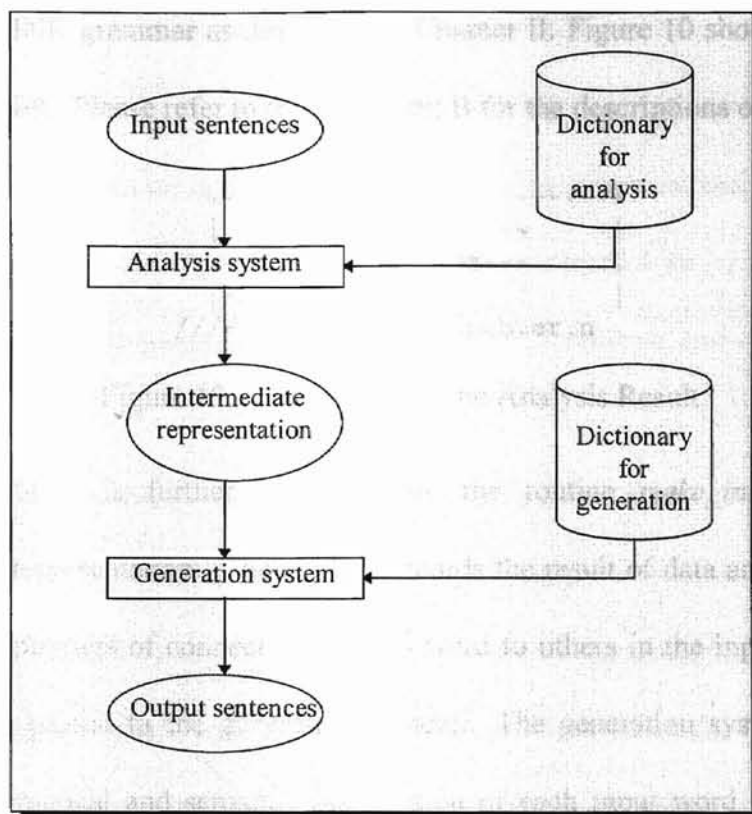


Figure 9. Overview of the Translation System

Operation

The input sentences are either typed in by the user or processed in batch mode. In a batch process, input sentences are written in a text file first, then the system processes the input file. The user specifically selects a domain (popular technology, politics, economy, or general information) which corresponds to the context of the input sentences. Although the system can handle two modes of translation, that is, manual and automatic, it is designed primarily to be an automatic translation system.

After the user feeds in the input sentences, control is passed to the analysis system. The analysis system produces a graphical representation of the input sentences,

based on the LINK grammar as described in Chapter II. Figure 10 shows an example of the analysis result. Please refer to the Appendix B for the descriptions of the connectors.

```

+-----Xp-----+
|               +---Ost---+
+-Wd-+-Ss+  +---Ds---+
|       |   |   |   |
///// she is my daughter.n .

```

Figure 10. An Example of the Analysis Result

This representation is further processed by the routine *make_inter* to create an 'intermediate' representation. The module records the result of data analysis in an array of words with pointers of connectors of each word to others in the input sentence. The control is then passed to the generation system. The generation system first gets the meaning, grammatical and semantic information of each input word by consulting the dictionary for the generation system. The generation dictionary uses a simple binary search tree to store up to ten different meanings for each English word. The meanings of a word are ranked based on the frequency of their use. In the case where an English word is not found in the generation dictionary, the system does a simple morphological analysis. This is carried out by looking at the ending of the unknown word. For example, if the ending of the unknown word is *ment* or *tion*, then it is highly possible that the unknown word is a noun. If the ending is *ize*, then the unknown word is given a verb transitive code.

Next, the generation system checks the type of the input sentence. After that, the system finds the main subject, predicate, and object of the sentence by scanning the input sentence from right-to-left. The nature of LINK grammar makes it easy to detect the main subject, predicate, and object of the input sentence. If a word has an S connector

pointing to its right, then it is assumed that the word is the subject of the sentence. A similar technique is employed to find the predicate of the input sentence by looking for a word with an O connector pointing to its right. The object of the input sentence is assumed to be those words lying on the right hand of a word with an O right connector. The system records the locations of the main subject, predicate and object of the input sentence.

After finding the predicted main subject, predicate and object of the input sentence, the system checks the part-of-speech of each word by looking at the result of the analysis system and by checking the data provided by the generation dictionary. If in the input sentence there is a word with more than one part-of-speech, the analysis system chooses the best part-of-speech and adds a specific ending to the word. For example, if the chosen part-of-speech of the word is a noun, then an ending '*n*' is added to the word. If it is an adjective, then an ending '*a*' is added. By looking at the ending of an input word, the generation system chooses the meaning for the word with more than one part-of-speech.

Next, the system finds possible idioms in the input sentence. The system does this by looking at the intermediate representation. In this system, an idiom is regarded as a word unit. If there is an ID or K connector, then words connected by those connectors are idioms. The system also checks the possibility of an idiom by checking the data provided by the dictionary for the generation system. This is carried out by looking at the idiom look-up table. The table lists the combinations of parts-of-speech that can lead to an idiom. If such combinations of parts-of-speech exist in the input sentence, the system concatenates the words and copies the result into a temporary new word and then consults

the generation dictionary to check whether the new word is in the dictionary or not. If it is, the system then erases the meaning of words that constitute the idiom except the first word. For instance, the meanings of the words *the*, *other*, and *hand* in the idiom *in the other hand* are erased. The meaning and other grammatical information of the idiom are assigned to the word *in*.

The next step is to check the type of the verbs to determine whether they are in a transitive, intransitive, passive, or active form or in a combination of the forms. Next, the system checks the part-of-speech of each word of the new input sentence (because an idiom or a collocation is detected). Checking part-of-speech by looking at the data provided by the dictionary for the generation system is rather complicated. The system checks the part-of-speech of adjacent words and then makes some changes to the words. There are a number of specific cases in changing the meaning of an input word. For instance, the meaning of an *a* or *an* is changed to *seekor*, *sebuah*, or *seorang* according to the part-of-speech of the following word. If the following word is a 'human noun' such as *teacher*, *student*, etc., then the meaning of *a* or *an* is changed to *seorang*. *Seekor* is used to change the meaning of an *a* or *an* if the following word is an 'animal noun' such as *horse* or *cat*. *Sebuah* is used for an abstract noun (*invitation*, *meeting*, etc.) or a concrete noun (*hotel*, *book*, etc.).

Sometimes the system needs to change the grammatical information of a word based on the data provided by the intermediate representation. The analysis system uses a certain connector type to relate a pair of words. To connect a verb with an adverb, the analysis system uses an E connector. If the meaning of the word connected by the E connector is not an adverb, then the system checks whether there is another meaning with

the part-of-speech of adverb or not. If there is one, then the system changes the previous meaning and its grammatical information. Next, the system changes the order of the words, e.g. from adjective-noun and possessive-noun sequence to noun-adjective and noun-possessive respectively. The next thing left to do is to concatenate the meaning of each input word to produce an output sentence. The output sentences are either shown in the terminal or saved to a file for post-editing purposes. If the user utilizes a batch process to translate texts, the results are written in a file named 'output'. The user can then do post-editing jobs on the file.

To solve the problem of homographs, the system uses the frequency of use of a meaning. A homograph is a word that has more than one meaning and the meanings have the same part-of-speech. An example of a homograph is the word '*bank*' which has at least two meanings (a financial institution or a part of a river). Every time a meaning of an English word is chosen, the frequency of use of the meaning is increased by one. In the case where there is a homograph, the system selects a meaning with the highest frequency of use. The user may change the result of the translation process by simply changing the frequency of use of a meaning. There is a simple dictionary editor program to do this.

There is a separate generation dictionary for each domain (general information, popular technology, economy, and politics). The number of words contained in the generation dictionary for the domains of general information, politics, economy and popular technology are 5,493 words, 5,344 words, 5,325 words, and 5,401 words respectively. The generation dictionary stores the English words in their full forms. The number of stem words for each domain is approximately 3,500.

System Performance

As stated earlier, articles taken from newspapers and journals are used to test the capability of the system. The study starts with the assumption that the articles taken are grammatically correct. However, the assumption is not always true because some of the sentences are not grammatically correct. This causes the analysis system to fail in analyzing some input sentences. Although it is common that input sentences are edited first before a machine translation system is invoked, this study leaves the input sentences unchanged in order to measure the actual performance of the system on natural text. The first article was taken from the *Inside Indonesia* journal edition number 53 (edition January - March 1998). The article was about politics and consisted of 87 sentences. The second article was taken from the *Asiaweek* magazine edition March 20, 1998. It was about economy and consisted of 24 sentences. The third article was taken from the *Asiaweek* magazine edition March 13, 1998. It was about popular technology and consisted of 55 sentences.

Before changes were made in the analysis system, the system could handle 64 sentences out of 87 sentences in the first article. For the second article, the analysis system could analyze 19 out of 24 sentences, while in the third article, the analysis system could parse 36 out of 55 sentences.

After some changes were made in the analysis system, the system could tackle 73 out of 87 sentences. In the second article, the analysis system could manage 23 out of 24 sentences while in the third article, the analysis system could analyze 44 out of 55 sentences. Table III gives a resume of the performance of the analysis system.

Table III.

Performance of the Analysis System

	Original (number of successfully analyzed sentences)	After Modification (number of successfully analyzed sentences)
Article 1	64 out of 87 sentences (73.56 %)	73 out of 87 sentences (83.91 %)
Article 2	19 out of 24 sentences (79.17 %)	23 out of 24 sentences (95.83 %)
Article 3	36 out of 55 sentences (65.55 %)	44 out of 55 sentences (80.00 %)

The following discussion briefly explains the techniques used to improve the performance of the analysis system. Appendix B contains the explanations of connectors in the following examples:

1. Adding new words and idioms into the analysis dictionary

The analysis system may fail to analyze the input sentence simply because there is an unknown word or idiom in the input sentence. An example of this occurs in the following sentence:

*Mehmet II delivered the **coup de grace** to the last significant remnant of the Roman Empire and bastion of the Eastern Orthodox Church by capturing Constantinople in 1453.*

Here, the idiom *coup de grace* is French, and by adding the idiom to the analysis dictionary, the analysis system can successfully parse the sentence.

2. Introducing a new connector type in the analysis rules

The rules of analysis system are complicated in the sense that changing a single rule is almost impossible because the rules are interconnected. Therefore, introducing a new type of connector for the linking requirements of a word is an alternative way to

improve the system's performance. The following sentence is an example of this case:

Even if unsuccessful, the attempt could trigger deep divisions within the elite, including the officer corps.

The modification made is for the word *if*.

The original rule for the word *if* is as follows:

if: (Cs+ & {Xc+ & {Xd-}} & (Wd- & (Qd+ or Ws+ or Wq+))) or ((Cs+ or [Mg+] or [Mv+]) & (({Xd- & Xc+} & MVs-) or ({Xc+ or {Xd-}} & CO+))) or (QI- & Cs+);

The rule is changed to the following:

if: (Cs+ & {Xc+ & {Xd-}} & (Wd- & (Qd+ or Ws+ or Wq+))) or ((Cs+ or [Mg+] or [Mv+]) & (({Xd- & Xc+} & MVs-) or ({Xc+ or {Xd-}} & CO+))) or (QI- & Cs+) or (E- & IFa+);

Here, the new connector is *IFa*, meaning that the word *if* can be connected to an adjective.

3. Changing the analysis rules

This is the most difficult technique used to make the performance of the analysis system better. Changing an analysis rule should be done very carefully because this affects the working of the system. Changing a rule might be good for a certain form of sentence but it might also cause the system to fail in analyzing other forms of sentences. The following sentence needs a change in the original analysis rule:

When you plan to turn your images into prints, use the highest resolution available for best quality.

Originally, the analysis rule for the word *best* is as follows:

best: ({Xc+} & {NR-} & {[[@Ec-]]} & L-) or ({NR- or ND-} & DD- & ((<noun-sub-x> & {TO+} & <noun-main-x>) or MVa-)) or [[E+]];

The rule is then modified to the following:

best: ({Xc+} & {NR-} & {[[@Ec-]]} & L-) or A+ or ({NR- or ND-} & DD- & ((<noun-sub-x> & {TON+} & <noun-main-x>) or MVa-)) or [[E+]];

Here, as we can see by comparing the two rules, the original rule for the word *best* does not imply that the word might serve as an adjective. Therefore, the analysis fails to recognize the sentence. By adding the linking requirement of an adjective for the word *best*, the analysis can parse the sentence.

Of the 73 sentences of the first article that can be analyzed, the generation system produces 61 'good' sentences (here, good means that the meaning of the output sentence is immediately clear), 10 'understandable' sentences (meaning that the reader has to think for a while before he/she can understand the meaning), and 2 'not good' sentences (meaning that the output sentences are not understandable). In the second article, of the 23 successfully analyzed sentences, the generation system generates 21 'good' sentences, 2 'understandable' sentences, and no 'not good' sentences. For the third article, of the 44 successfully parsed input sentences, the generation system gives 40 'good' sentences, 3 'understandable' sentences, and one 'not good' sentence. Table IV gives a summary of the performance of the generation system where the calculations presented are based on the work of the analysis system. Table V presents the overall performance of the system. In the Table V, numbers in the column 'Not Good' represent the number of the 'not understandable' output sentences plus the number of unanalyzed input sentences.

Table IV.

Performance of the Generation System

(Calculations Based on the Results of the Analysis System)

	Good (need no thinking)	Understandable (need thinking)	Not Good (not understandable)
Article 1	61 out of 73 sentences (83.56 %)	10 out of 73 sentences (13.70 %)	2 out of 73 sentences (2.74 %)
Article 2	21 out of 23 sentences (91.30 %)	2 out of 23 sentences (8.69 %)	0 (0 %)
Article 3	40 out of 44 sentences (90.91 %)	3 out of 44 sentences (6.82 %)	1 out of 44 sentences (2.27 %)

Table V.

The Overall Performance of the System

	Good (need no thinking)	Understandable (need thinking)	Not Good (not understandable and unanalyzed sentences)
Article 1	61 out of 87 sentences (70.11 %)	10 out of 87 sentences (11.50 %)	16 out of 87 sentences (18.39 %)
Article 2	21 out of 24 sentences (87.50 %)	2 out of 24 sentences (8.33 %)	1 out of 24 sentences (4.17 %)
Article 3	40 out of 55 sentences (72.73 %)	3 out of 55 sentences (5.45 %)	12 out of 55 sentences (21.82 %)

The system translates an input sentence in a relatively short time, typically in the range of 0.2 - 250 seconds. Most sentences are processed less than 5 seconds. The system needs around 1 minute to process the second article (about economy and consists of 24 sentences), approximately 7 minutes to process the first article (about politics and

comprised of 87 sentences), and about 6 minutes to process the third article (about popular technology and consists of 55 sentences). It should be noted here that there are two long sentences in the first article and one long sentence in the third article which are skipped in the translation process. This is because the analysis system takes a long time (hours) to parse the sentences. It is thus assumed that the analysis system fails in processing all three sentences. The two long sentences in the first article are as follows:

Suharto's unseating of Megawati Sukarnoputri from the leadership of the Indonesian Democratic Party, and the party's virtual destruction in the lead-up to the May 1997 parliamentary elections, crippled an avenue of opposition to Suharto's reelection in March 1998 and stripped away the veneer of democracy Pancasila doctrine espouses.

Suharto's recent authorisation of government regulations providing for five star generals within Abri and his own promotion to five star general, along with the late Sudirman and Nasution, can be seen as a move to remind Abri officers, in the lead-up to the March 1998 elections, that he is supreme commander of Abri.

The one long sentence of the third article is as follows:

You'll want: Macromedia's Flash or Shockwave (www.macromedia.com), animation programs that are widely used by Web content authors; Real Networks' Real Player (www.real.com), a mainstay for listening to music samples and sound bites, and for viewing the occasional video clip; Apple's Quick Time (www.quicktime.apple.com), a multimedia format used on many news sites; Adobe's Acrobat Reader (www.adobe.com), which allows you to download, view and print documents such as detailed charts and graphs.

Sometimes there is a word that – because of the context of the sentence – does not have the proper Indonesian meaning. The word *co-opting* in the following sentence is an example of the case.

*Suharto's new-found devotion to Islam has been useful in **co-opting** Islam and for providing a counter-foil to Abri, much as the PKI did for Sukarno.*

Here, the meaning of *co-opting* is '*to adopt the ideas of Islam and use them for political ends*'. Since there is no proper translation for the word, its meaning is assigned to mean '*to win*'.

CHAPTER V
EVALUATION, CONCLUSIONS, AND SUGGESTED
FUTURE RESEARCH

Evaluation

The LINK grammar system is good for analyzing and recognizing a grammatical English sentence. The generation system is 'robust' in the sense that it produces an output sentence whatever the result of the analysis system is. Many output sentences are unnatural, that is, they are somewhat difficult to read. However, a native Indonesian speaker can comprehend the meaning of these 'strange' output sentences. Some post-editing is needed to correct the unnatural output texts.

The system can not handle relations between sentences which are conveyed by the use of pronouns, definite articles, etc. The following example demonstrates this problem.

The problem was very difficult. We took hours to solve it.

In the above sentence, the word *it* refers to the word *problem*. However, the system can not make a correct interpretation of the word *it*. To minimize this problem, the system assigns the meaning of the word *it* as '[*hal itu*]' which roughly means "that matter". Unlike English and other languages, the Indonesian language does not have gender pronouns. To deal with this problem, the meanings of the English words 'he' and 'she' are assigned to '*dia{m}*' and '*dia{f}*' respectively where '{*m*}' means 'male' and '{*f}*' means 'female'.

Conclusions

This system obviously is not ready to be used for fully-automated machine translation, but it does reflect the general direction in which such a system might be designed. However, the system presented in this thesis can handle many phenomena found in the English language such as noun-verb agreement, questions, imperatives, complex and irregular verbs, punctuation, commas, prepositions, adverbs, relative clauses, possessives, and many other things.

Although it is first assumed that sentences of magazine and journal articles are written carefully and therefore grammatically correct, the study finds that the assumption is not always true. Some of the sentences are not grammatically correct and carefully written.

Suggested Future Research

One of the things that can be done to improve this system is to further categorize the part-of-speech of a word based on the semantic information it carries. For instance, nouns can be further classified as 'noun event', 'noun action', 'noun animate', 'noun inanimate', etc. Adding more semantic primitives to the information stored in the generation dictionary can help in choosing the correct meaning for a homograph, and therefore improve the performance of the system.

Another area of improvement to the system occurs with the addition of morphological analysis. The system uses a simple morphological analysis in deciding the part-of-speech of an unknown word by looking at the ending of the word. Morphological

analysis has the advantage of reducing the size of the dictionary by eliminating the need to store all forms of a word. This process can extract the stem of a word so that only the stem needs to be stored in the dictionary as in the case of a verb. A routine to check spelling can be used during morphological analysis to correct misspelled words so that the user does not have to submit the sentence again. The addition of morphological analysis can degrade the response time of the system, but as the size of the dictionary continues to increase, the reduced memory usage may be worth the extra processing time. Because many output sentences are strange to read, incorporating a word processor into the system would be helpful for the user to perform post-editing tasks.

SELECTED BIBLIOGRAPHY

- ALPAC. (1966). *Language and Machine: Computers in Translation and Linguistics*. Washington D.C: National Academy of Sciences.
- Appelo, L. and Landsbergen, J. (1986). The Machine Translation Project Rosetta. In Gerhardt (Ed.). *I. International Conference on the State of the Art in Machine Translation in America, Asia and Europe: Proceedings of IAI-MT-86*. (pp. 34-51). Saarbrücken: IAI/Eurotra-D.
- Arnold, D., Balkan, L. Humphreys, R. L., Meijer, S., and Sadler, L. (1994). *Machine Translation: An Introductory Guide*. Cambridge: NCC Blackwell.
- Bennet, W.S., and Slocum, J. (1985). The LRC Machine Translation System. *Computational Linguistics*, Vol. 11, 111-121.
- Chandioux, J. (1976). Météo: Un Système Opérationnel Pour La Traduction Automatique Des Bulletins Météorologiques Destinés Au Grand Public. *Meta* Vol. 21, 127-133.
- Grinberg, D., Lafferty, J., and Sleator, D. (1995). *A Robust Parsing Algorithm for LINK Grammar*. Technical Report CMU-CS-95-125. Pittsburgh, PA: Department of Computer Science, Carnegie Mellon University.
- Hutchins, W. J. (1986). *Machine Translation: Past, Present, Future*. England: Ellis Horwood Limited.
- Hutchins, W. J. and Somers, H. L. (1992). *An Introduction to Machine Translation*. New York: Academic Press.
- Josselson, H.H. (1972). Fourteen year summary of the work of the Wayne State University Machine translation project. In *Fourteenth annual report on research in computer-aided translation: Russian-English*. (pp. 124-140). Detroit: Wayne State University.
- Kwee, J.B. (1993). *Teach Yourself Indonesian*. Chicago: NTC Publishing Group.
- Maas, H.-D. (1977). The Saarbrücken Automatic Translation System. In *Third European Congress on Information System and Networks*. Munich: Verlag Dokumentation.
- Maas, H.-D. (1987). The MT System Susy. In King, M. (Ed.). *Machine Translation Today: The State of The Art*. (pp. 209-246). Edinburgh University Press.
- MacDonald, R. R., and Darjowidjojo, S. (1967). *Indonesian Reference Grammar*. Washington D.C.: Georgetown University Press.

- Melcuk, I. A. (1988). *Dependency Syntax: Theory and Practice*. New York: State University of New York Press.
- Nagao, M. (1989). *Machine Translation: How Far Can It Go?*. New York: Oxford University Press.
- Sleator, D. and Temperley, D. (1991). *Parsing English with a LINK Grammar*. Technical Report CMU-CS-91-196. Pittsburgh, PA.: Department of Computer Science, Carnegie Mellon University.
- Slocum, J. (1987). METAL: The LRC Machine Translation System. In King, M. (Ed.). *Machine Translation Today: The State of The Art*. (pp. 319-350). Edinburgh University Press.
- Toma, P. (1976). An Operational Machine Translation System. In Brislin, R.W. (Ed.). *Translation: applications and research*. (pp. 247-259). New York: Gardner Press.
- Toma, P. (1977). *Systran as A Multilingual Machine Translation System*. In Third European Congress on Information System and Networks. (pp. 569-581). Munich: Verlag Dokumentation.
- Vasconcellos, M. (Ed.). (1988). *Technology as Translation Strategy*. (American Translators Association Scholarly Monograph Series, Volume II), State University of New York at Binghamton.
- Vauquois, B. (1975). *La Traduction automatique à Grenoble*. Paris: Dunod.
- Weaver, W. (1949). Translation. in Locke, W.N. and Booth, A.D. (Eds.). *Machine Translation of Languages* (pp. 15-23). New York: Technology Press of MIT and Wiley and Sons.
- Witkam, A.P.M. (1983). *Distributed Language Translation: Feasibility Study of A Multilingual Facility for Videotex Information Networks*. Utrecht: BSO.

GENERAL REFERENCES

- Aho, A.V., and Ullmand, J. D. (1972). *The Theory of Parsing, Translation, and Compiling*. Englewood Cliffs, NJ: Prentice-Hall.
- Allen, J. (1987). *Natural Language Understanding*. California: The Benjamin/Cummings Publishing Company.
- Gazdar, G and Mellish, C. (1989). *Natural Language Processing in LISP: An Introduction to Computational Linguistics*. Workingham, England: Addison-Wesley Publishing Company.
- King, M. (ed.). (1983). *Parsing Natural Language*. London: Academic Press Inc.
- King, M. (ed.). (1983). *Machine Translation Today: The State of The Art*. Edinburgh: Edinburgh University Press.
- Newton, J. (ed.). (1992). *Computers in Translation: A Practical Appraisal*. New York: Routledge.
- Nirenburg, S. (ed.). (1992). *Machine Translation: Theoretical and Methodological Issues*. England: Cambridge University Press.
- Whitelock, P. and Kilby, K. (1995). *Linguistic and Computational Techniques in Machine Translation System Design*. London: UCL Press.
- Schank, R. C. and Colby, K. M. (ed.). (1973). *Computer Models of Thought and Language*. San Fransisco: W. H. Freeman and Company.

APPENDICES

APPENDIX A
USERS'S MANUAL

MOHAMMAD TEDUH ULINIANSYAH
 USER'S GUIDE FOR TRANS AND GDICT
 JULY 23, 1998

PURPOSE: Trans is an English-to-Indonesian machine translation system designed to translate English sentences into Indonesian sentences.

The system has two modes of translation:

1. Manual mode

In this translation mode, user should type the following command in the user prompt.

```
% trans analysis_dictionary_file domain_name
```

The name of the *analysis_dictionary_file* is *ana_dict*.

The available *domain_name* now are: *general*, *politic*, *economy*, and *technology*.

After the translation prompt appear, the user can then type in the input sentence.

2. Batch translation mode

In this translation mode, the user may create a file with one sentence on each line.

To make a comment in the input file, the user should precede the line with an '%'. Figure 11 gives an example of an input file for a batch translation mode.

The user then should type the following command in the user prompt

```
% trans analysis_dictionary_file domain_name < input_file > output_file
```

The available *domain_name* now are: *general*, *politic*, *economy*, and *technology*.

The system will create a file named '*output*' which consists of the output sentences. The user then may perform post-editing on this file.

```
%AsiaWeek, March 20, 1998
Among the lessons to be learned from East Asia's financial turmoil, one
stands out: the need for more monitoring.
The rapid and unfettered accumulation of short-term debt was one major
factor that contributed to last year's meltdown.
This has prompted calls for heightened surveillance, including a wider role
for the Bank for International Settlements (BIS), the Swiss-based central
banks' central bank.
"One of the ways the governments of Asia know how much short-term debt
they have is not from their own accounts, but from the BIS," says Harvard
University economist Jeffrey Sachs.
```

Figure 11. A Sample Input File

PURPOSE: Gdict is an utility program to see and edit the contents of the generation dictionary.

In the user prompt, the user may type the following command to activate gdict:

```
% gdict domain_name
```

A user menu will appear. The user may choose an action according to the menu available where he/she can add, delete, or find a word. The user may also change the meaning and grammatical information of an English word.

APPENDIX B
DESCRIPTIONS OF CONNECTORS

A: connects pre-noun ("attributive") adjectives to following nouns.

CO: used to connect 'openers' to subject of clauses.

Cs: connects conjunctions and certain verbs with the subject of clauses.

DD: connects definite determiners ("the", "his") to number expressions certain things like
number expressions and adjectives acting as nouns.

Ds: links a determiner and a noun.

E: links a verb-modifying adverbs which precede the verb.

IFa: connects the word 'if' with an adjective.

L: connects certain determiners to superlative adjectives.

Mg: connects noun with present participles.

MVa: connects adverbs to preceding verbs or adjectives.

MVs: connects verbs and adjectives to modifying phrases like adverbs, prepositional
phrases, time expressions, certain conjunctions, "than"-phrases, and other things.

Mv: connects noun with passive participles.

ND: connects numbers with certain expressions which require numerical determiners.

NR: connects fraction words with superlatives.

<noun-main-...>: a macro; the "main" connectors for nouns, used to link them to the rest
of the sentence (as subject, object, etc.).

<noun-sub-...>: a macro; the "sub" connectors for nouns, used to link them to modifiers
like prepositional phrases and relative clauses.

Ost: links a 'to be' verb with its objects.

Qd: used in questions, 'd' means declarative.

QI: connects certain verbs and adjectives to question-words, forming indirect questions.

Ss: connects subject-nouns to finite verbs, 's' means singular.

TO_n: used with nouns that take infinitival complements.

W_d: used to express an ordinary declarative sentences. It connects the main clause to the

WALL ("/////"). The WALL is a dummy word at the beginning of every input sentence and is automatically inserted by the analysis system.

W_q: used in object questions.

W_s: used in subject questions.

X_c: used to connect a word to a coma to the right.

X_d: connects a word to a comma to the left.

X_p: used to connect the wall with the punctuation in the end of the sentence.

APPENDIX C
SOURCE AND TRANSLATED TEXTS

Article 1

Source: Inside Indonesia Journal, edition 53. January - March 1998

1. Senior Abri intentions may be obscure now, says Bob Lowry, but no one should assume they will remain that way.
Maksud Abri senior mungkin tak jelas sekarang, Bob Lowry berkata, tetapi tak seorangpun harus menganggap bahwa mereka akan tetap seperti itu .
2. In 1451 Mehmet II replaced his father, Murad II, as Sultan of the Ottoman Turks.
Di 1451 Mehmet II mengganti ayah dia{m} , Murad II , sebagai Sultan dari Turks Ottoman .
3. He was the third son, by a slave girl, from the Sultan's harem.
Dia{m} adalah anak laki-laki ketiga , oleh seorang wanita budak , dari haremnnya Sultan .
4. Murad's first two sons, by high born wives, were expected to be his successors but the eldest died of natural causes and the second was murdered in suspicious circumstances.
Pertamanya Murad dua anak laki-laki , oleh istri-istri kalangan atas , diharapkan menjadi pengganti dia{m} tetapi yang tertua [tersebut] meninggal dari sebab-sebab alami dan kedua [tersebut] dibunuh di keadaan mencurigakan .
5. Among his first acts on assuming power, Mehmet II had his infant half brother drowned in his bath.
Diantara aksi pertama dia{m} dalam mengambil kekuasaan, Mehmet II menyuruh saudara tiri laki-laki muda dia{m} ditenggelamkan di mandi dia{m} .
6. Fratricide was common place following dynastic successions and was designed to reduce the field of possible usurpers and deny disgruntled citizens a symbol of opposition.
Pembunuhan saudara kandung adalah mengikuti biasa dinasti suksesi dan didesain untuk mengurangi kesempatan dari perebut kekuasaan mungkin dan menyangkal warga tak puas sebuah simbol dari oposisi .
7. Mehmet II delivered the coup de grace to the last significant remnant of the Roman Empire and bastion of the Eastern Orthodox Church by capturing Constantinople in 1453.
Mehmet II mengirim kudeta ke terakhir sisa-sisa signifikan dari Kekaisaran Romawi dan benteng pertahanan dari Gereja Ortodoks Timur dengan merebut Constantinople di 1453 .
8. The empire was no more but the Orthodox Church survives in Russia and Eastern Europe.
Kerajaan tidak ada lagi tetapi Gereja Ortodoks selamat di Rusia dan Eropa Timur .

9. Interestingly, in the six years before his death, Sultan Murad attempted to retire, handing power to Mehmet II.
Secara menarik, di enam tahun sebelum kematian dia{m} , Sultan Murad mencoba untuk mengundurkan diri , memberi kekuasaan ke Mehmet II .
10. But he was eventually called back because of Mehmet's youth, and dissension among his advisers.
Tetapi dia{m} akhirnya dipanggil kembali karena dari kemudaannya Mehmet , dan perselisihan diantara penasehat hukum dia{m} .
11. Nevertheless, Mehmet became one of the greatest leaders of the Ottoman Empire.
Meskipun demikian, Mehmet menjadi satu dari pemimpin yang paling besar dari Empire Ottoman .
12. This historical cameo reminds us of the inherent difficulties authoritarian regimes experience replacing leaders and establishing and maintaining their legitimacy.
Cameo bersejarah ini mengingatkan kami/kita dari kesulitan-kesulitan khas rejim otoriter mengalami mengganti pemimpin dan mendirikan dan memelihara legitimasi mereka .
13. Although fratricide is now unfashionable and much has changed since the days of Mehmet II, the nature of the human animal remains constant.
Meskipun pembunuhan saudara kandung sekarang tidak mode dan banyak telah dirubah sejak jaman dari Mehmet II, sifat dari kebinatangan manusia tinggal/tetap konstan .
14. Whoever replaces Suharto will need the backing of influential sections of the elite, especially those in the armed forces.
Siapapun yang mengganti Suharto akan butuh dukungan dari seksi berpengaruh dari golongan elit [tersebut] , khususnya itu di golongan militer .
15. It might seem that Suharto has promoted a tame bunch of professional lackeys to the senior leadership of Abri.
[hal itu] kelihatannya bahwa Suharto telah mempromosikan sebuah kumpulan jinak dari antek-antek profesional ke kepemimpinan senior dari Abri .
16. But to expect all of them to remain docile when the redistribution of power and privilege is on offer is to ignore history and deny human nature.
Analysis failed
17. Suharto's actions in relation to Abri show that he is acutely aware of this problem.
Tindakannya Suharto di relasi ke Abri tunjukkan bahwa dia{m} benar-benar sadar akan dari masalah ini .
18. From the beginning of his reign he has been a master of divide and rule tactics and the use of political eunuchs.

Dari permulaan dari pemerintahan dia{m} dia{m} seorang ahli dari siasat 'pecah-belah dan kuasai' dan pemakaian [tersebut] dari praktek pengebirian politik .

19. He successfully isolated and sidelined potential challengers like Generals Nasution, Sumitro, Murtopo, and Jusuf among others.
Dia{m} dengan sukses mengucilkan dan sisihkan penantang berpotensi seperti Jendral-jendral Nasution , Sumitro , Murtopo , dan Jusuf diantara yang lain .
20. Abri command arrangements also make it difficult for any commander to act in isolation or in concert with fellow malcontents.
Perencanaan komando Abri juga membuat [hal itu] susah untuk komandan untuk beraksi di [aksi]pengisolasian atau kesesuaian dengan teman tak puas .
21. The intelligence system is pervasive and comprises several independent parts.
Sistim intel adalah dapat ditembus dan terdiri dari beberapa bagian bebas .
22. And the capital Jakarta is secured by units from several different commands, reducing the risk of Praetorian arrogance.
Dan ibu kota Jakarta yang dilindungi oleh unit dari beberapa komando berbeda , mengurangi risiko dari kesombongan Praetorian .
23. Overlaying the structure of power are personnel arrangements designed to minimise the opportunity for officers to betray their political masters.
Analysis failed
24. Many of the senior commanders are personally known to the president, having been personal aides, presidential security officers, or local commanders.
Banyak dari komandan senior secara personal diketahui ke Presiden , pernah menjadi ajudan pribadi , petugas keamanan presiden , atau komandan lokal .
25. Most are also tied to the president by patronage, either directly or indirectly.
Kebanyakan juga diikat ke Presiden oleh perlindungan , apakah secara langsung atau secara tidak langsung .
26. Finally, most senior military appointments are regularly rotated.
Akhirnya, kebanyakan pengangkatan militer senior secara reguler diputar .
27. This has the effect of denying officers the opportunity to develop independent power bases and alliances.
Ini mempunyai efek meniadakan bagi petugas kesempatan [tersebut] untuk mengembangkan dasar-dasar kekuasaan bebas dan persekutuan .
28. With a retirement age of 55 and a top heavy structure it also ensures rapid promotion for favoured officers, flattering their egos and creating a sense of obligation to their patrons.
Analysis failed

29. If Abri has been so effectively bent to the president's will can it play any significant role in the presidential succession?
Kalau Abri begitu secara efektif cenderung untuk kehendaknya Presiden dapatkah [hal itu] memainkan peran signifikan di suksesi presiden ?
30. The answer is that it will depend on the circumstances at the time.
Jawaban bahwa [hal itu] akan tergantung pada keadaan di waktu [tersebut] .
31. Abri officers might act to save the state as defined by doctrine, from self-interest, or from a combination of both.
Analysis failed
32. Abri has contingency plans to cover possibilities like the death of the incumbent president and the forthcoming special session of the People's Consultative Assembly (MPR).
Abri mempunyai rencana darurat ke kemungkinan-kemungkinan tutup seperti kematian dari Presiden yang sedang memangku jabatan dan sesion khusus yang akan datang dari Majelis Permusyawaratan Rakyat [tersebut] (MPR) .
33. These plans only cover security matters, they cannot address the political consequences except in a procedural manner.
Rencana ini hanya mencakup masalah keamanan , mereka tidak dapat mengalamatkan konsekwensi politik kecuali di sebuah cara prosedural .
34. Nevertheless, they will put some senior commanders in potentially advantageous positions when implemented.
Meskipun demikian, mereka akan menaruh beberapa komandan senior di secara potensial posisi-posisi menguntungkan ketika dilaksanakan .
35. Ironically, the regime's own indoctrination is a double edged sword.
Secara ironis, indoktrinasinya rejim adalah pedang bermata dua .
36. Doctrine states that Abri is above politics, that it is the guardian of the nation and the protector of the people from false prophets.
Analysis failed
37. Were the succession to give rise to unrest, for whatever reason, some officers might feel duty bound to save the state from those who would destroy the regime.
Analysis failed
38. Likewise with President Suharto's policy of maintaining dynamic tension and competition among senior officers.
Sebagaimana dengan Presiden Suharto kebijaksanaan memelihara ketegangan dinamik dan kompetisi diantara petugas senior .

39. Although this assists in maintaining his own position it could lead to factionalisation during a succession.

Meskipun ini membantu dalam memelihara posisi dia{m} sendiri [hal itu] dapat menuju ke faksionalisasi pada waktu sebuah suksesi .

40. The Suharto succession will occur as a consequence of the president's death or infirmity, or a decision to step down either at a regular MPR session or mid-term.

Suksesi Suharto akan terjadi sebagai sebuah konsekwensi dari Presiden [tersebut] kematian atau kelemahan , atau sebuah keputusan untuk mengundurkan diri apakah di sesion MPR reguler atau mid-term .

41. Death or infirmity or retirement mid-term would be the easiest solution if the incumbent vice president was generally acceptable and competent and the process was not drawn out.

Kematian atau kelemahan atau mid-term pengunduran diri akan menjadi pemecahan yang paling mudah kalau wakil presiden yang sedang memangku jabatan secara umum dapat diterima dan berkompeten dan proses tidak berlarut-larut .

42. A prolonged succession, as a result of creeping senility or of a decision to step down at a regular MPR session, poses the greatest challenge.

Sebuah suksesi lama , sebagai sebuah hasil dari usia tua yang merangkak atau dari sebuah keputusan untuk mengundurkan diri di sesion MPR reguler , menghadapi tantangan yang paling besar .

43. It would dissolve patronage obligations and create an opportunity for a number of challengers to emerge.

[hal itu] akan membubarkan kewajiban perlindungan dan membuat sebuah kesempatan untuk sejumlah dari penantang untuk muncul .

44. In the absence of strong political institutions it is possible that challengers could seek endorsement directly from the populace, through social and religious organisations, or from the military.

Di ketidakhadiran dari institusi politik kuat [hal itu] adalah mungkin bahwa penantang dapat mencari secara langsung pengesahan dari rakyat , melalui sosial dan agamis organisasi , atau dari militer [tersebut] .

45. Such endeavours could divide the loyalty of Abri commanders and staff, forcing officers to declare their position.

Usaha seperti itu dapat membagi kesetiaan dari komandan Abri dan staf , memaksa petugas untuk menyatakan posisi mereka .

46. When personal ambition and other motives such as envy, revenge and frustration are added, the potential for splits within Abri increases.

Ketika ambisi pribadi dan motif yang lain seperti kecemburuan, pembalasan dendam dan kekecewaan tambah, potensi untuk perpecahan dalam Abri bertambah .

47. Should civil unrest arise it will inevitably put more power into the hands of the troops deployed to restore order.
Seandainya kegelisahan sipil bangkit[hal itu] akan tak terelakkan menaruh lebih kekuasaan ke tangan dari serdadu yang disebarkan untuk memulihkan keadaan .
48. Likewise with the option of semi-retirement, the so-called Lee Kuan Yew option.
Analysis failed
49. Indonesia lacks the political infrastructure provided by Singapore's People's Action Party.
Indonesia kekurangan infrastruktur politik disediakan oleh Partai Aksi Rakyatnya Singapura .
50. After more than 30 years of the New Order, Abri remains the repository of Suharto's power.
Sesudah lebih daripada bertahun-tahun 30 dari Order Baru , Abri tinggal/tetap tempat penyimpanan dari kekuasaannya Suharto .
51. And, as Sultan Murad found, any attempt to divide power at the peak allows scope for divisive politics from below.
Dan, sebagai Sultan Murad menemukan, usaha untuk membagi kekuasaan di puncak [tersebut] mengijinkan kesempatan untuk politik yang bersifat memecah-belah dari bawah .
52. Any attempt to explore the lines of cleavage in Abri is confounded by the nature of the regime and restrictions on research.
Analysis failed
53. The Indonesian officer corps is sometimes defined according to patronage links, divisional allegiances, military classes, professional orientation, religious affiliations, and ethnicity.
Petugas Indonesia korps kadang-kadang didefinisikan menurut hubungan perlindungan , divisi kesetiaan , kelas militer , orientasi profesional , afiliasi agamis , dan ke-etnisan .
54. None of these, in themselves, are of much utility because interests and loyalties change with time and circumstance.
Analysis failed
55. Rather than explore these imponderables, differences of interest can be explored in relation to concrete issues like the succession and political transition.
Daripada menjelajahi faktor-faktor yang tak terhitung ini , perbedaan-perbedaan dari kepentingan dapat dijelajahi di relasi ke isu kongkrit seperti suksesi dan politik politik transisi .
56. The officer corps both active and retired is split on the succession issue.

Korps petugas yang aktif dan mengundurkan diri dipecah di isu suksesi .

57. Many believe that it is long past time that Suharto retired but few are prepared to say so directly or openly.
Banyak percaya bahwa[hal itu] adalah waktu masa lalu panjang bahwa Suharto mengundurkan diri tetapi sedikit adalah siap untuk berkata begitu secara langsung atau secara terbuka .
58. There are many reasons for their reticence.
Ada banyak alasan untuk sikap bungkam diri mereka .
59. Suharto can still severely penalise critics directly or indirectly, for example by denying passports, business or educational opportunities to the individual or his family.
Suharto dapat masih secara berat menghukum kritikus secara langsung atau secara tidak langsung , contohnya dengan menyangkal paspor , bisnis atau kesempatan-kesempatan pendidikan ke individu atau keluarga dia{m} .
60. They are also reluctant to publicly acknowledge divisions within the Abri family for fear of stimulating cleavages which might be exploited by other opponents of the regime.
Mereka juga enggan untuk secara luas mengakui pembagian dalam keluarga Abri akan ketakutan untuk menstimulasi perpecahan yang mungkin dieksploitasi oleh lawan yang lain dari rejim [tersebut] .
61. The regime also fosters subservience to authority and deference to age.
Rejim juga membantu sikap mengabdikan ke penguasa dan rasa hormat ke usia .
62. The motives for opposing Suharto's continued rule also vary.
Motif untuk menentang berkelanjutannya peraturan Suharto juga bervariasi .
63. Some resent being left out of power and patronage, others resent the privileges given to family and friends of the president.
Analysis failed
64. Some resent Suharto's usurpation of the New Order, some believe that he has done a good job but that his time has passed.
Beberapa membenci pengambil-alihan kekuasaannya Suharto dari Order Baru , beberapa percaya bahwa dia{m} telah kerjakan sebuah pekerjaan bagus tetapi bahwa waktu dia{m} telah lewat .
65. Others are critical of the means he has used to retain power.
Yang lain adalah kritis dari alat dia{m} telah memakai untuk tetap menguasai kekuasaan .

66. Suharto's unseating of Megawati Sukarnoputri from the leadership of the Indonesian Democratic Party, and the party's virtual destruction in the lead-up to the May 1997 parliamentary elections, crippled an avenue of opposition to Suharto's reelection in March 1998 and stripped away the veneer of democracy Pancasila doctrine espouses.
Analysis failed
67. Suharto's manipulation of Islam for political ends also causes anxiety among some members of the officer corps.
Manipulasinya Suharto dari Islam untuk tujuan-tujuan politik juga menyebabkan kekhawatiran diantara beberapa anggota dari petugas [tersebut] korps .
68. Suharto's new-found devotion to Islam has been useful in co-opting Islam and for providing a counter-foil to Abri, much as the PKI did for Sukarno.
Suharto baru ketaatan ke Islam berguna dalam memenangkan Islam dan untuk menyediakan sebuah batu sandungan ke Abri , sebagaimana PKI [tersebut] melakukan untuk Sukarno .
69. But many officers, including the current minister for defence and security, fear that the political mobilisation of Islam could rebound on the regime.
Tetapi banyak petugas, termasuk menteri saat ini untuk pertahanan dan keamanan, takut bahwa mobilisasi politik dari Islam dapat mengganjal rejim [tersebut] .
70. They fear the emotive mobilizing power of Islam in the wrong hands.
Mereka takut kekuasaan memobilisasi emosional dari Islam di tangan yang salah .
71. Political transition is the other issue on which divisions within Abri are apparent.
Transisi politik adalah isu yang lain dimana yang pembagian dalam Abri adalah nyata .
72. For some, political transition means retention of the current regime.
Untuk beberapa, transisi politik berarti penyimpanan dari rejim saat ini .
73. For others it includes liberalisation of the current regime.
Untuk yang lain[hal itu] termasuk proses liberalisasi dari rejim saat ini .
74. For a minority, mainly of the younger generation, it includes eventual transition to liberal democracy.
Untuk sebuah golongan minoritas , terutama dari generasi lebih muda ,[hal itu] termasuk transisi akhir ke demokrasi bebas .
75. In the cut and thrust of politics the two issues sometimes become entwined.
Di pemotongan dan serangan serangan dari politik kedua isu [tersebut] kadang-kadang menjadi dijalin .
76. Opposition to Suharto is mainly expressed by retired officers or the occasional maverick in Abri's parliamentary faction.

Oposisi ke Suharto terutama dinyatakan oleh petugas yang pensiun atau organisasi informal berkala di fraksi parlementernya Abri .

77. Nevertheless, they reflect the views of many serving officers who do not have the latitude to express their views.
Meskipun demikian, mereka merefleksikan pandangan dari banyak melayani petugas yang tidak memiliki kebebasan [tersebut] untuk menyatakan pandangan mereka .
78. A serving officer risks his career prospects and consideration for sinecures on retirement by openly opposing Suharto's continued rule.
Seorang petugas yang sedang bertugas mempertaruhkan prospek karir dia{m} dan pertimbangan untuk jabatan yang kering di pengunduran diri dengan secara terbuka menentang berkelanjutannya peraturan Suharto .
79. Suharto's recent authorisation of government regulations providing for five star generals within Abri and his own promotion to five star general, along with the late Sudirman and Nasution, can be seen as a move to remind Abri officers, in the lead-up to the March 1998 elections, that he is supreme commander of Abri.
Analysis failed
80. That Suharto needs these symbols of power shows that he recognizes that he is losing power as age catches up with him and discussion of succession mounts.
Bahwa Suharto butuh simbol ini dari pertunjukan kekuasaan bahwa dia{m} mengenali bahwa dia{m} kehilangan kekuasaan sebagai usia mengejar dia{m} dan diskusi dari suksesi memuncak .
81. The discussion thus far is centred on elite politics, but a revolt by more junior officers inspired by ideological zeal or religious zealotry cannot be ignored.
Diskusi sampai saat ini difokuskan di politik elit , tetapi sebuah pemberontakan oleh lebih petugas junior yang diinspirasi oleh semangat ideologis atau fanatisme agamis tidak dapat diabaikan .
82. For example, the currency crisis related to the floating of the rupiah, the devastating fires that swept Indonesia recently and increasing concerns about corruption and the abuse of power could motivate such action.
Contohnya , krisis mata uang berhubungan ke pengembangan dari rupiah , kebakaran yang menghancurkan yang menyapu akhir-akhir ini Indonesia dan bertambah perhatian tentang korupsi dan penyalahgunaan dari kekuasaan dapat memotivasi tindakan seperti itu .
83. Even if unsuccessful, the attempt could trigger deep divisions within the elite, including the officer corps.
Bahkan jika tak sukses, usaha dapat memicu pembagian dalam dalam golongan elit , termasuk petugas korps .

84. In summary, there is little to be gained from looking for individual military officers or fixed constellations of officers who would be king or kingmakers.
Analysis failed
85. The commanders in place when Suharto decides to retire or is no more will only show their true colours when the time comes.
Analysis failed
86. What colours they fly will depend on their latitude for action and the character of the individuals concerned.
Warna yang mereka kemukakan akan tergantung pada kebebasan mereka untuk tindakan dan karakter dari individu dimenyangkut . ?
87. Those who would delete Abri from the analysis of presidential succession would do well to reflect on the nature of the human animal.
Mereka yang akan menghapus Abri dari analisa suksesi presiden akan melakukan dengan baik untuk merefleksikan di sifat dari kebinatangan manusia .

Article 2

Source: Asiaweek Magazine, edition March 20, 1998

1. Among the lessons to be learned from East Asia's financial turmoil, one stands out: the need for more monitoring.
Diantara pelajaran menjadi dipelajari dari kerusuhan keuangannya Asia Timur , satu menonjol : kebutuhan untuk lebih memonitor .
2. The rapid and unfettered accumulation of short-term debt was one major factor that contributed to last year's meltdown.
Akumulasi yang cepat dan tak terkekang dari hutang jangka pendek satu faktor utama yang menyumbang ke kekacauannya tahun yang lalu .
3. This has prompted calls for heightened surveillance, including a wider role for the Bank for International Settlements (BIS), the Swiss-based central banks' central bank.
Ini telah mendorong panggilan untuk pengawasan yang diperkuat , termasuk sebuah peran lebih luas untuk Bank untuk Persengketaan Internasional (BIS) , bank-bank sentral berbasis pada Swiss nya bank sentral .
4. "One of the ways the governments of Asia know how much short-term debt they have is not from their own accounts, but from the BIS," says Harvard University economist Jeffrey Sachs.
Satu dari jalan bahwa pemerintah dari Asia tahu berapa hutang jangka pendek mereka memiliki bukan dari rekening mereka sendiri , kecuali dari BIS [tersebut] , ahli ekonomi universitas Harvard Jeffrey Sachs berkata .

5. "I've always believed that the BIS should be expanded to cover the emerging markets more aggressively."
Saya telah selalu percaya bahwa BIS seharusnya diperluas untuk lebih mencakup secara agresif pasar yang baru .
6. Set up in 1930, the BIS started out as an exclusive club of mainly Western central banks.
Didirikan di 1930, BIS mulai sebagai sebuah grup eksklusif dari sebagian besar bank-bank sentral Barat .
7. It has operated as a clearing house for the foreign reserves of many countries and as a lender of short-term loans.
[hal itu] telah dioperasikan sebagai sebuah kantor kliring untuk cadangan asing dari banyak negara-negara dan sebagai peminjam dari pinjaman jangka pendek .
8. Since 1996, membership has expanded from 32 to include nine more monetary authorities in Asia, Latin America and Europe.
Sejak 1996, keanggotaan telah diperluas dari 32 untuk memasukkan sembilan lebih ahli-ahli moneter di Asia , Amerika Latin dan Eropa .
9. The Asian economies now in the group are China, Hong Kong, India, Japan, Singapore and South Korea.
Analysis failed
10. Owned and controlled by the member central banks, the BIS provides them with data on capital flows and sets recommended regulatory standards such as capital-adequacy levels.
Dipunyai dan dikontrol oleh bank-bank sentral anggota , BIS menyediakan mereka dengan data di aliran modal dan standar pengaturan menganjurkan kumpulan seperti level kecukupan modal .
11. Recently, the BIS has been helping central banks with short-term liquidity problems, aiding institutions in Latin America, Eastern Europe and Asia with crucial bridging loans.
Akhir-akhir ini, BIS menolong bank-bank sentral dengan jangka pendek masalah likuiditas , institusi yang membantu di Amerika Latin , Eropa Timur dan Asia dengan pinjaman bantuan penting .
12. In August last year, the BIS arranged a \$3.3-billion facility for Thailand.
Di bulan Agustus tahun yang lalu , BIS mengatur fasilitas 3.3-billion \$ untuk Thailand .
13. Partly because of perceptions that the International Monetary Fund has become too powerful, governments are now considering a wider role for the BIS.

Sebagian karena dari persepsi bahwa Dana Moneter Internasional (IMF) telah menjadi sangat kuat, pemerintah sedang sekarang mempertimbangkan sebuah peran lebih luas untuk BIS [tersebut] .

14. Pakistan Finance Minister Sartaj Aziz says that the BIS could provide daily reports on the trading levels of each currency to give monetary authorities a better grasp of currency market movements.
Pakistan Menteri Keuangan Sartaj Aziz berkata bahwa BIS dapat menyediakan laporan perhari di level perdagangan dari setiap mata uang untuk memberi ahli-ahli moneter genggaman lebih baik dari gerakan pasar mata uang .
15. Some blamed speculative currency trading for East Asia's financial meltdown.
Beberapa menyalahkan perdagangan perdagangan mata uang spekulatif untuk kekacauan keuangannya Asia Timur .
16. Mark Hansen, Jakarta-based banking expert with consulting firm Booz-Allen & Hamilton, says the BIS could be instrumental in upgrading banking standards and providing technical assistance to regulators.
Mark Hansen, ahli perbankan berbasis di Jakarta dengan konsultan Booz-Allen & Hamilton, berkata bahwa BIS dapat menjadi penolong dalam memperbaiki perbankan standar dan menyediakan bantuan teknis ke pemerintah .
17. There have also been calls for the creation of an Asian BIS to enhance regional cooperation in such areas as bank supervision and reserve management.
Ada juga panggilan untuk penciptaan dari BIS Asia untuk mempertinggi kerjasama regional di daerah seperti itu seperti pengawasan bank dan pengelolaan cadangan .
18. A group of 11 East Asian central banks already meets informally to discuss these issues.
Sebuah kelompok dari 11 bank-bank sentral Asia Asia Timur sudah bertemu secara informal untuk mendiskusikan isu ini .
19. Philippine Central Bank Governor Gabriel Singson, a key promoter of an Asian BIS, says that the idea needs to be discussed further in light of the economic crisis.
Gubernur Bank Sentral Filipina Gabriel Singson , seorang promotor kunci dari BIS Asia , berkata bahwa ide butuh menjadi mendiskusikan selanjutnya mengenai dari krisis ekonomi .
20. One essential question is whether the bank's role would go beyond monitoring to providing financial assistance.
Satu pertanyaan penting adalah apakah perannya bank akan pergi melebihi memonitor daripada menyediakan bantuan keuangan .
21. If so, member banks would have to be prepared to contribute funds for that purpose.

Kalau begitu, bank-bank anggota akan menjadi siap untuk menyumbang dana untuk tujuan itu .

22. "Asia must find a solution to the debt market," says Hong Kong Financial Secretary Donald Tsang Yam-kuen.
Asia harus temukan sebuah pemecahan ke institusi keuangan , Hong Kong Sekretaris Keuangan Donald Tsang Yam-kuen berkata .
23. "We have no instrument to meet short-term balance-of-payments difficulties."
Kami/kita tidak memiliki instrumen untuk memenuhi kesulitan-kesulitan neraca pembayaran jangka pendek .
24. A beefed-up BIS or an Asian-based version could fill that role.
Sebuah BIS yang diperkuat atau sebuah versi berbasis pada Asia dapat mengisi peran itu .

Article 3

Source: Asiaweek Magazine, edition March 13, 1998

1. I recently purchased my first digital camera.
Saya akhir-akhir ini membeli kamera digital pertama saya .
2. I'm pretty happy with it, but when I make prints the photos don't seem to be very sharp.
Saya agak gembira dengan [hal itu] , tetapi ketika saya membuat cetakan foto tidak kelihatannya menjadi sangat tajam .
3. I have a good printer.
Saya memiliki sebuah printer bagus .
4. Am I doing something wrong?
Apakah saya kerjakan sesuatu salah ?
5. Although they are rapidly improving, prints from digital cameras costing less than \$1,000 cannot yet match the quality of film.
Meskipun mereka sedang dengan cepat dimembaik, cetakan dari berharga kamera digital kurang daripada \$ 1,000 tidak dapat masih menandingi kualitas dari film .
6. But there are some techniques you can use to improve results.
Tetapi ada beberapa teknik kamu dapat pakai untuk memperbaiki hasil-hasil .
7. Many digital cameras allow you to select the resolution of each shot -- the crispness and detail of the image.

Banyak kamera digital memungkinkan kamu untuk memilih resolusi dari setiap bidikan - ketajaman dan detil detil dari gambar [tersebut] .

8. It is expressed in terms of the number of pixels -- basic picture elements -- arranged in a grid.
[hal itu] dinyatakan dalam dari nomor dari pixels -- elemen gambar mendasar -- diatur di grid .
9. Cameras today commonly offer a standard resolution of 640 by 480 pixels, and some also allow lower-resolution shots of 320 by 240 pixels.
Kamera hari ini biasa menawarkan sebuah resolusi standar dari 640 kali pixels 480 , dan beberapa juga memungkinkan bidikan beresolusi lebih rendah dari 320 kali 240 pixels .
10. When you plan to turn your images into prints, use the highest resolution available for best quality.
Ketika kamu merencanakan untuk untuk merubah gambar kamu ke cetakan , memakai resolusi yang paling tinggi tersedia untuk kualitas terbaik .
11. But remember, higher-resolution images occupy more of the camera's precious memory, cutting the total number of shots you can take.
Analysis failed
12. Lower-resolution images generally look okay displayed on a computer, so go that route if you don't want hard copies.
Analysis failed
13. Another tip: unlike ordinary photos, digital shots don't enlarge well.
Analysis failed
14. Individual picture elements, when blown up, start to look like tiny squares.
Elemen gambar individu , ketika diperbesar , mulai untuk terlihat seperti seperti empat persegi kecil .
15. The phenomenon is called pixelization, and if you've played older video games you've seen the characteristic blocky objects and jagged borders before.
Fenomena dipanggil pixelization , dan kalau kamu telah memainkan permainan video lebih tua kamu telah melihat obyek seperti kotak khas dan batas bergerigi sebelumnya .
16. This means digital shots taken at too great a distance cannot be salvaged in the darkroom by enlarging the subject to fill the frame.
Analysis failed

17. For better images, make sure that what you are shooting looms large in the viewfinder or monitor.
Analysis failed
18. You might want to take several shots, each time moving in a little closer, to make sure you get something usable.
Analysis failed
19. If your camera has a review feature, you can then select the best of the batch and delete the rest to free up memory.
Kalau kamera kamu mempunyai sebuah fasilitas pengulangan , kamu dapat lalu memilih terbaik dari batch [tersebut] dan menghapus sisanya untuk membebaskan memori .
20. I notice DVD-ROM drives are now available for PCs.
Saya memperhatikan bahwa penggerak DVD-ROM sekarang tersedia untuk komputer personal .
21. Should I buy one?
Haruskah saya membeli satu ?
22. Purchasing the latest hardware can be risky -- no one wants to be stuck with a Betamax.
Membeli perangkat keras terakhir dapat berisiko-- tak seorangpun ingin terikat dengan Betamax .
23. When deciding, it helps to ask: Is the flashy new widget going to become standard equipment?
Ketika memutuskan, [hal itu] menolong untuk bertanya: apakah widget baru menyilaukan akan menjadi untuk peralatan standar ?
24. Is there a wide variety of software available for it?
Apakah sebuah macam lebar dari perangkat lunak tersedia untuk [hal itu] ?
25. Are prices falling rapidly?
Apakah harga sedang jatuh dengan cepat ?
26. Most industry analysts agree that within the next five years, DVD (short for digital video disk or digital versatile disk) will replace CD-ROM as a high-capacity data storage system for PCs.
Analysis failed
27. But the transition has only just begun, and patience might be a virtue.
Tetapi transisi sudah hanya baru saja mulai , dan kesabaran mungkin sebuah kebaikan .

28. Content, in the form of movies, is now widely available on DVD.
Isi, dalam bentuk dari film, sekarang secara umum tersedia di DVD .
29. But PC game manufacturers and software publishers have yet to make the switch, so DVD-ROM is not yet a desktop must-have.
Tetapi produsen permainan komputer personal dan penerbit perangkat lunak masih belum perlu untuk membuat pergantian [tersebut] , sehingga DVD-ROM bukan masih sebuah harus memiliki desktop .
30. Drive prices have plunged from \$1,000 to around \$500 in the past 12 months or so, and should continue to fall.
Analysis failed
31. A nagging compatibility issue clouds the future.
Isu kesesuaian berlanjut-lanjut memuramkan masa depan [tersebut] .
32. Due on the market next year are DVD systems that not only play disks, but also record data -- for example, TV programs, home videos and the contents of your hard drive.
Due di pasar tahun berikutnya adalah sistim DVD yang tidak hanya hanya memainkan disks, tetapi juga rekam data-- contohnya , program televisi , video rumah dan isi [tersebut] dari penggerak hard disk kamu . ?
33. Alas, manufacturers are still at war over a standard format.
Susahnya, produsen masih dalam peperangan tentang sebuah format standar .
34. It probably won't be clear soon whether a disk recorded on a DVD drive of the future will be able to play in a DVD-ROM drive bought today.
[hal itu] mungkin tidak akan jelas segera apakah disk direkam di sebuah penggerak DVD dari masa depan [tersebut] akan sanggup untuk memainkan di sebuah penggerak DVD-ROM dibeli hari ini .
35. On the Internet, I sometimes come across a Web page with a box that looks like it has missing graphics.
Di Internet [tersebut] , saya kadang-kadang menemukan sebuah halaman Web dengan kotak yang terlihat seperti [hal itu] mempunyai grafik yang hilang .
36. When I click on it, a message pops up to tell me I need a "plug-in."
Ketika saya mengklik di[hal itu], sebuah pesan muncul untuk mengatakan pada saya bahwa saya butuh sebuah plug-in .
37. Can you explain what that is and why plug-ins are necessary?
Dapatkan kamu menerangkan yang itu dan mengapa plug-in adalah perlu ?

38. In the beginning, the average Internet browser was like a cheap Swiss Army Knife: it had a few crude implements, all of them rather dull.
Pada awalnya [tersebut], browser Internet rata-rata seperti Knife Swiss Army murah :[hal itu] mempunyai beberapa yang sederhana melaksanakan, semua dari mereka agak tumpul.
39. Browsers could display text coded in HTML, the universal programming language used to create basic Web pages, as well as simple graphics.
Browser dapat mempertunjukkan teks ditulis di HTML, bahasa pemrograman universal memakai untuk membuat halaman Web mendasar, seperti juga grafik yang mudah.
40. But to experience some of the cooler aspects of the Web (audio, video and animation, for example), you needed -- still need, for that matter -- additional software blades.
Analysis failed
41. That's where plug-ins come in.
Itu dimana plug-in datang.
42. Sometimes known as applets, they are small programs that add to your browser's capabilities, allowing it to perform tasks beyond ordinary page display.
Kadang-kadang diketahui sebagai applets, mereka adalah program kecil yang menambah ke kemampuannya browser kamu, mengizinkan [hal itu] untuk menampilkan tugas melebihi pameran halaman biasa.
43. Plug-ins are usually available for free and can be downloaded directly from the Internet.
Plug-in biasanya tersedia dengan gratis dan dapat diambil secara langsung dari Internet.
44. Unfortunately there's a very high signal-to-noise ratio (see Geekspeak below) surrounding plug-ins.
Analysis failed
45. It seems that every novice programmer with a home page is creating content that can be sampled only with software created specifically for that purpose.
[hal itu] kelihatannya bahwa setiap ahli pemrograman pemula dengan sebuah home pej sedang membuat isi yang yang dapat dicoba hanya dengan perangkat lunak dibuat secara spesifik untuk tujuan itu.
46. There are literally hundreds of plug-ins available.
Ada benar-benar ratusan dari plug-in tersedia.
47. But you don't need to clutter up your hard drive unduly.

Tetapi kamu tidak butuh untuk mengacaukan dengan sangat penggerak hard disk kamu .

48. There are a handful of essential ones.
Ada beberapa dari yang penting .
49. You'll want: Macromedia's Flash or Shockwave (www.macromedia.com), animation programs that are widely used by Web content authors; Real Networks' Real Player (www.real.com), a mainstay for listening to music samples and sound bites, and for viewing the occasional video clip; Apple's Quick Time (www.quicktime.apple.com), a multimedia format used on many news sites; Adobe's Acrobat Reader (www.adobe.com), which allows you to download, view and print documents such as detailed charts and graphs.
Analysis failed
50. Acrobat is popular on personal finance sites and in online catalogs.
Acrobat adalah populer di tempat biaya pribadi dan di katalog online .
51. A hot new plug-in is Beatnik (www.headspace.com), which plays high-fidelity music from the Internet.
Plug-in populer dan baru Beatnik(www.headspace.com), yang memainkan musik berkualitas baik dari Internet [tersebut] .
52. The latest browsers from Microsoft and Netscape are supposed to download plug-ins more or less automatically when you come across a site that requires them.
Terakhir browser dari Microsoft dan Netscape seharusnya untuk mengambil plug-in kurang lebih secara otomatis ketika kamu menemukan sebuah tempat yang memerlukan mereka .
53. Often you will have to do the downloading yourself.
Seringkali kamu akan harus untuk melakukan pengambilan [tersebut] sendiri .
54. To check out the plug-ins available, visit www.microsoft.com or www.netscape.com.
Untuk memeriksa plug-in tersedia , kunjungan www.microsoft.com atau www.netscape.com .
55. Or you can go to CNet's www.download.com.
Atau kamu dapat pergi ke www.download.comnya CNet .

VITA

Mohammad Teduh Uliniansyah

Candidate for the Degree of

Master of Science

Thesis: **AN ENGLISH-TO-INDONESIAN MACHINE TRANSLATION SYSTEM**

Major Field: Computer Science

Biographical:

Personal Data: Born in Banjarmasin, Kalimantan Selatan, Indonesia, on April 9, 1967, the son of Mohammad Ansar Soedirman and Yetti Ansar.

Education: Graduated from SMAN 34 High School, Jakarta, Indonesia, in 1985; received Bachelor of Science degree in Telecommunication Engineering from The Institute of Technology of Shibaura, Tokyo, Japan in April 1991.

Completed the requirements for the Master of Science degree with a major in Computer Science at Oklahoma State University in July 1998.

Experience: Researcher Staff, The Agency for the Assessment and Application of Technology, Jakarta, Indonesia, since May, 1991. Developer of the Indonesian Generation System at the Multilingual Machine Translation System (MMTS) project, 1991 to 1995.