MACHINE TRANSLATION OF NOUN PHRASES: FROM ENGLISH TO ARABIC

by
Azza Abd El-Moniem Mohamed

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE
in

Computer Engineering

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2000

MACHINE TRANSLATION OF NOUN PHRASES: FROM ENGLISH TO ARABIC

By
Azza Abd El-Moniem Mohamed

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
Computer Engineering

Under the Supervision of

Prof. Dr. Ahmed Rafea
Faculty of Computers & Information
Computer Science Dept.
Cairo University

Prof. Dr. Hoda Baraka
Faculty of Engineering
Computer Engineering Dept.
Cairo University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2000

MACHINE TRANSLATION OF NOUN PHRASES: FROM ENGLISH TO ARABIC

by

Azza Abd El-Moniem Mohamed

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in

Computer Engineering

Approved by the
Examining Committee:

Prof. Dr. Hoda Anis Barka

____

Prof. Dr. Ahmed Abd El-Wahed Rafea

____

Prof. Dr. Ali Hassan Fahmy

____

Prof. Dr. Ayman El-Dessouky Ibrahim

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2000

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation and gratitude to the supervisors' committee Prof. Dr. Ahmad Rafea and Prof. Dr. Hoda Baraka for their contribution to the thesis.

I would like to express my deepest appreciation to Prof. Dr. Ahmad Rafea for suggesting the research topic, for his continuous guidance, for his patience and for enriching me with his precious and valuable knowledge and experience.

I would like to express my deepest gratitude to Prof. Dr. Hoda Baraka for her cooperation, for her continuous support and for her vast experience.

Last but not least I would like to extend my high appreciation to Dr. Khaled Shaalan (my dear Husband) and for his encouragement, his valuable advises, and for the fruitful discussions that we have had together. Furthermore I would like to thanks Yara (my daughter) and Ahmed (my son) for their consideration.

Deep thanks to all my colleagues in Central Laboratory for Agricultural Expert System (CLAES) at Agricultural Research Center, Egyptian Ministry of Agriculture and Land Reclamation.

ABSTRACT

In the modern world, there is an increased need for language translation. Attempts of language translation are as old as computer themselves. Machine translation is the attempt to automate all, or part of the process of translating from one human language to another.

The present work reports our attempt in developing a transfer-based machine translation system of English noun phrases into Arabic. English noun phrases are frequently used in scientific and technical documents. According to transfer approach of machine translation, the system consists of three main modules, responsible for analysis, transfer, and generation. The analysis component assigns grammatical structures to the input noun phrase by means of English grammatical rules and an English monolingual dictionary. The transfer component builds target language equivalents of the source language grammatical structures by means of a comparative grammar that relates every source language representation to some corresponding target language representation. This involves a bilingual dictionary. The generation component provides the target language translation. This involves the synthesis grammar rules of Arabic and an Arabic monolingual dictionary.

A major design goal of this system is that it can be used as a stand-alone tool and can be very well integrated with a general machine translation system for English sentence. The system is implemented in Prolog and the parser is written in DCG formalism. Experiment on real noun phrases was performed. The thesis also described our experience with the developed machine translation system and reports results of its application on real titles of thesis and journals.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Attempts at computer translation of human languages are as old as computer themselves. This chapter will briefly sketch some background on natural language processing and machine translation. Then the aim of the current work and the structure of this thesis are summarized.

1.1 Overview

The study of natural language has been an important area of artificial intelligence almost since the beginning of the field. Two main goals motivate AI work on natural language. One is the theoretical goal, and close to that of the linguist, namely, to discover how we use language to communicate. The other is technological goal, namely, to enable the intelligent computer interfaces of the future, where natural language becomes an important means for man-machine interaction. Luckily, progress toward one of these goals often is progress toward the other-a better theoretical understanding leads to more robust systems, and a better understanding of processing issues in actual applications suggests new goals and techniques of theoretical interest. The ultimate solution to language understanding must wait until we can effectively model almost all aspects of human intelligence. Many applications, however, do not require full conversational capabilities or encyclopedic knowledge. For instance, a natural language interface that serves as a query language to a database need only focus on questions and can limit the language it understands to concepts that arise in the database. On the contrary, machine translation (MT) is the attempt to automate all, or part of the process of translating from one human language to another. It is one of the oldest large-scale applications of computer science. In today's increasingly networked world, the need for systems to translate documents to and from a variety of languages is expanding, for applications as diverse as:

- Multilingual e-mail
- Browsing (such as on the World Wide Web) texts in other languages
- High-quality translation of business letters and reports
- Translation of technical documents and articles
- Speech-to-speech translation for business and travel.

While useful MT technology is currently available, it is not yet capable of providing both high-quality and wide-domain performance simultaneously. For higher quality, the domain may be limited, and human assistance required while for wider domain, output quality may be sacrificed. MT research continues to push the boundaries of this automation-quality-scope continuum.

1.2 What is machine translation?

People who need documents translated often ask themselves whether they could use a computer to do the job. When a computer translates an entire document automatically and then presents it to a human, the process is called machine translation (*MT*). When a human composes a translation, perhaps calling on a computer for assistance in specific tasks such as looking up specialized words and expressions in a dictionary, the process is called human translation. There is a gray area between human and machine translation, in which the computer may retrieve whole sentences of previously translated text and make minor adjustments as needed. However, even in this gray area, each sentence was originally the result of either human translation or machine translation. We will reserve the label "machine translation" for the case when a computer performs both the initial translations of the sentences and subsequent manipulations. All else we will call "translator tools".

1.3 The Translation Tripod

Machine translation is highly appealing when its quality is acceptable for some purpose with little or no human revision. For machine translation to be appropriate, it must be sitting on a stable "tripod" Such as the figure 1-1.

A translation project can be thought of as sitting on a tripod whose three legs (Kay, 1997) are the source text, the specifications, and the terminology. If any of the three legs is removed, the project falls down.

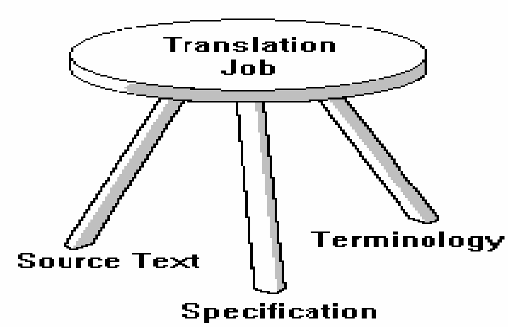


Figure 1-1 the components of translation project

Source text

Obviously, no translation can be done without a source text (i.e., the document to be translated). But for machine translation, an additional basic requirement is that the source text be available in machine-readable form. That is, it must come on diskette or cartridge or tape or by modem and end up as a text file on your disk. A fax of the source text is not considered to be in machine-readable form, even if it is in a computer file. A fax in a computer file is only a graphical image of the text, and the computer does not know which dots compose the letter a or the letter b. Conversion of a source text on paper or in a graphical image file to machine-readable form using imaged character recognition (ICR) is not usually accurate enough to be used without human editing, and human editing is expensive, adding an unacceptable cost component to the total cost of machine translation. Thus, for machine translation to be appropriate, it is usually necessary to obtain the word processing or desktop publishing file from the organization that created the source text. But this is only one of many requirements.

Specifications

All translation projects have specifications. The problem is that they are seldom written down. Specifications tell how the source text is to be translated. One specification that is always given is what language to translate into. But that is insufficient. Should the format of the target text (i.e., the translation) be the same as that of the source text or different? Who is the intended audience for the target text? Does the level of language need to be adjusted? In technical translation, perhaps the most important specification is what equivalents to use for technical terms. Are there other target texts with which this translation should be consistent? What is the purpose of the translation? If the purpose is just to get a general idea of the content of the source text, then the specifications would include "indicative translation only." An indicative translation is usually for the benefit of one person rather than for publication and need not be a high-quality translation. Thus, publication-quality translations are high-quality translations (and are usually the result of human translation), while indicative translations are low-quality translations (and are usually the result of machine translation). These two types of translation are not normally in competition with each other, since a requester of translation will typically want one type or the other for a given document and a given set of specifications. Sometimes, the two types are complementary, such as when an indicative translation is used to decide whether or not to request a high-quality translation of a particular document. In this environment, an indicative translation may be requested for a number of documents, and, using the indicative translations, the requester may select one or two documents for publication quality translation. As previously mentioned, indicative translations are usually done using machine translation and high-quality translations are usually done using human translation. This fact reveals a basic difference between humans and computers. Humans, with proper study and practice, are good at producing high-quality translations but typically can only translate a few hundred words an hour to approximately a thousand words an hour, depending on such factors as the difficulty of the source text. Even with very familiar material, how fast they can type or dictate their translations limits human translators. Computers are good at producing low-quality translations very quickly. Some machine translation systems can translate tens of thousands of words an hour. But as they are "trained" by adding to their dictionaries and grammars, they reach a plateau where the quality of the output does not improve. By upgrading to a more powerful computer, the speed of translation improves but not the quality. By upgrading to a "more powerful" human translator, the quality of translation improves but not necessarily the speed. Here we have a classic case of a trade-off. You can have high speed or high quality but not both. Indicative translation (high speed, low cost, but low quality) represents a new and growing market but does not substantially overlap with the existing market for publication quality translation. If, on the more likely hand, your specifications include high-quality translation, then it is not obvious that machine translation is appropriate for your current translation job. Here quality would be measured by whether the target text is grammatical, accurate, understandable, readable, and usable. Usability can be measured by selecting tasks, such as maintenance operations, which can be accomplished by a source language reader with the help of the source text and seeing whether those same tasks can be performed by a target-language reader with the help of the target text. Such measurements are notoriously expensive, but a skilled reviewer can accurately predict usability simply by studying the source and target texts. A target-language monolingual person can measure grammaticality, and understandability, and readability, which are progressively more stringent requirements. But accuracy requires the assistance of a skilled bilingual person who examines both the source and target texts.

Terminology

The treatment of terminology could have been included soley under specifications. But terminology is so important that the actual terminological database (also called a "termbase") supplied with a source text has been listed as a third essential component of a translation job. The aspect of terminology that does fit under specifications is the requirement that the translation job use a certain termbase into order to achieve desired consistency. Let us explain what we mean by consistency. Translation requesters typically want the terminology in their translated documents to mesh closely with terminology in related documents. For example, a software company will want all revisions of a software manual to use the same terms as the original, to avoid confusing readers. Translation requesters should track all terminology relevant to a given document and deliver that terminology to the translation provider along with specifications and source text. The specification component of the job tells what appropriate termbase to use and, as is all too common, tells what to do if a source-text term is missing from the termbase. The terminology component of the job contains the termbase itself. Now we can define an appropriate translation job (for a human or for a computer) as one that sits on a stable tripod. It must include a source text (in machine-readable form if for machine translation); it must include well-defined follow the specifications; and it must include any specified termbase. In addition, we can define an appropriate translation as a translation that combines the source text and the termbase in a way that matches the specifications. Note that we said "appropriate" translation, not "good" translation. A poor (low-quality) translation may be appropriate if the specifications include a requirement for a fast, indicative translation.

1.4 Why Machine Translation Is Hard

Many factors contribute to the difficulty of machine translation, including words with multiple meanings, sentences with multiple grammatical structures, uncertainty about what a pronoun refers to, and other problems of grammar. But two common misunderstandings make translation seem altogether simpler than it is. First, translation is not primarily a linguistic operation, and second, translation is not an operation that preserves meaning. There is a famous old example that makes the first point well. Consider the sentence:

The police refused the students a permit because they feared violence.

Suppose that it is to be translated into a language like Arabic in which the word for 'police' ('الشرطة ') is feminine. Presumably the pronoun that translates 'they' will also have to be feminine. Now replace the word 'feared' with 'advocated'. Now, suddenly, it seems that 'they' refers to the students and not to the police and, if the word for students ('الطلاب ') is masculine, it will therefore require a different translation. The knowledge required to reach these conclusions has nothing linguistic about it. It has to do with everyday facts about students, police, violence, and the kinds of relationships we have seen these things enter into. The second point is, of course, closely related. Consider the following question, "Where do you want me to put myself?" but it is a very natural translation for a whole family of English questions of the form "Where do you want me to sit/stand/sign my name/park/tie up my boat?" In most situations, the English "Where do you want me?" would be acceptable, but it is natural and routine to add or delete information in order to produce a fluent translation. Sometimes it cannot be avoided because there are languages like French, as well as Arabic, In which pronouns must show number and gender, Japanese where pronouns are often omitted altogether, Russian where there are no articles, Chinese where nouns do not differentiate singular and plural nor verbs present and past, and German where flexibility of the word order can leave uncertainties about what is the subject and what is the object.

1.5 The Traditional Structure of Machine Translation Systems

While there have been many variants, most MT systems, and certainly those that have found practical application, have parts that can be named for the chapters in a linguistic textbook. They have lexical, morphological, syntactic, and possibly semantic components, one for each of the two languages, for treating basic words, complex words, sentences and meanings. Each feeds into the next until the last one in the chain produces a very abstract representation of the sentence. There is also a 'transfer' component, the only one that is specialized for a particular pair of languages, which converts the most abstract source representation that can be achieved into a corresponding abstract target representation. The target sentence is produced from this essentially by reversing the analysis process. Some systems make use of a so-called 'interlingua' or intermediate language, in which case the transfer stage is divided into two steps, one translating a source sentence into the interlingua and the other translating the result of this into an abstract representation in the target language.

1.6 Natural language processing

Natural language processing (NLP) can be defined, in a very general way, as the discipline having as its ultimate, very ambitious goal that of enabling people to interact with machines using their "natural" faculties and skills. This means, in practice, that machines should be able to understand spoken or written sentences constructed according to the rules of some natural language, and should be capable of generating in reply meaningful sentence in this language.

The task of NLP is that of accepting inputs in a human natural language, and to transform the inputs into some sort of formal statements that are to be "meaningful" for a computer. The computer will be, therefore, able to react correctly to the given input; sometimes, the reaction will take the form of a NL "answer," i.e., the computer will use the formal representation corresponding to the analysis of the input to generate, in turn, statements in natural language. NLP is characterized by the presence of some, very primitive and idiosyncratic indeed, form of "understanding" of the "meaning" of a given statement. As a consequence, we will exclude from the description of the NLP domain some trivial and purely passive forms of processing of NL inputs. Examples are the simple transfer on magnetic support of a spoken input through the use of a voice recorder, or the handling of inputs formed by single words, e.g., all sort commands, entered by a keyboard, or spoken through a voice recognition system. We consider, in fact that a real problem of "meaning" begins only when several words combine together inside a written string or an utterance.

1.7 The standard paradigm for NLP

While there have been many variants, the structure of most MT systems have parts that can be named for the chapters in a linguistic textbook. They have lexical, morphological, syntactic, and possibly semantic components, one for each of the two languages, for treating basic words, complex words, sentences and meanings. Each feeds into the next until the last one in the chain produces a very abstract representation of the sentence.

1.7.1 Source language analysis

Natural language analysis is the process of mapping between a natural language text and a representation of its form and/or content. This representation can be a syntactic structure representation, a representation of the text's prepositional meaning, a comprehensive interlingua text (consisting of unambiguous semantic propositions and discourse/pragmatic information) or some specialized representation geared at a particular application. In knowledge-based machine translation, the analysis stage is expected to produce a complete interlingua text. In essence, the quality of the translation depends up on the depth and quality of the analysis. Most transfer-based MT systems stop at a syntactic representation often augmented with semantic markers (such as case markers for verb arguments), although the trend is toward ever-deeper semantic analysis.

A comprehensive system of natural language analysis, such as an analysis module of a knowledge-based machine translation system, must include the following basic components:

- **Morphological analysis:** the decomposition of words into their uninflected root forms, performed at the word level. There are many morphological phenomena: almost all language has inflectional morphology; the majority has some form of derivational morphology. A number of general models of morphological processing have been investigated. At the theoretical level, the most popular approach to morphology is the so-called two-level approach (Koskenniemi 1983; Karttunen 1983). In practical systems many other, less general and more language and task-specific approaches have been used.
- **Syntactic analysis:** the extraction of all well-formed syntactic structures and dependencies for a source text, performed at the sentence level. In the MT environment, a grammar must be written for each source language, in one of the many current grammar formalisms, such as, for instance, Lexical Functional Grammar, Generalized Phrase Structure Grammar, Head-driven Phrase Structure Grammar, Definite Clause Grammar, Tree-adjoining Grammar or Government-and-Binding-related Grammars. The use of a "canonical" formalism facilitates the use of a single grammar interpreter applicable to any language whose grammar is defined in the selected formalism.
- **Semantic analysis:** the creation of the knowledge structures in a text-meaning representation language (interlingua in MT) that reflect the meanings of lexical units in the source text and semantic dependencies among them, performed at the sentence level but often having to take into account suprasentential

contexts. Semantic analysis procedures are typically developed for a particular domain (e.g. medicine, finance, and computers), though general, “common sense” semantic knowledge is also used. The existence of canonical formalisms for encoding world knowledge and text meaning enables the use of a single universal semantic interpreter with different knowledge source for each domain.

- **Pragmatic or discourse analysis:** suprasentential analysis leading to the resolution of anaphors, ellided phrases, deixis, as well as the attribution of intent and speech acts. In its full form, discourse analysis leads to the creation of a text-meaning structure in a representation language with the various domain-oriented and rhetorical relations among the elements of a text, including coreference of noun phrases and anaphors, causal and temporal relations, topic/comment structure and so forth. The state of the art in pragmatic and discourse analysis is not as well developed as the other three phases of language analysis.

1.7.2 Target language generation

The process of natural language generation, in its unconstrained form, starts with the specification of the "need to communicate," the prepositional goals for a target language text, and a pragmatic profile of the speech situation-knowledge about speaker/author (or, more generally, text producer), the hearer/reader (text consumer), the style of communication and so on. A generator then must perform the following tasks:

- **Content delimitation:** The system must select which of the active prepositional and rhetorical goals should be overtly realized in text and which should be left for the human consumer to infer.
- **Text structuring:** The system must determine the order of propositions and the boundaries of sentences in the target language text.
- **Lexical selection:** The system must select open-class lexical units to be used in the target language text.
- **Syntactic selection:** The system must select syntactic structures for the target language clauses and perform closed-class lexical selection according to syntactic structure decision.
- **Constituent ordering:** The system must establish the order of syntactic constituents in a sentence.
- **Realization:** the system must map from syntactic representation with lexical insertions into surface string.

1.8 Aim of the work

The aim of the current work is to develop a machine translation system, which translates an English noun phrase into Arabic. This system is based on the transfer approach that includes an English noun phrase analysis module, a transfer module, and a generation module. The system uses a bilingual dictionary, an English lexicon, and an Arabic Lexicon. The research is targeted at translating computer science titles of theses and journals. The MT system is being developed using Prolog language. A major design goal of this system is that it can be used as a stand-alone tool and can be very well integrated with a general MT system for English sentence.

1.9 Structure of the thesis

This thesis is organized as follows. Chapter 2 is a review on the field of the machine translation, it discusses the history and the different approaches of the machine translation. In Chapter 3 we propose the English-Arabic transfer based machine translation system, how it work, how we prepare our output to be evaluated by it and how this system is implemented. Chapter 4 presents the evaluation methodology used to evaluate our system. Chapter 5 concludes the thesis and gives directions for future research.

CHAPTER 2

MACHINE TRANSLATION

In exploring the field of MT, it helps to look separately at different related fields. In Section 2.1, we give the breakdowns of automatic translation that covers many activities. Section 2.2 presents a brief history of MT. In Section 2.3, we look inside the most common approaches to the non-human component of MT, so-called the translation engine. Section 2.4 concludes the chapter by comparing the transfer-based method, which is the main concern of this study, with the most related ones.

2.1 Basic Concepts

We would like to take the term *machine assisted translation* (MAT) as covering all techniques for automating the translation activity. The term *human-aided machine translation* (HAMT) should be reserved for the techniques, which rely on a real automation of the translating function with some human intervention in preedition, postedition, or interaction. The term *machine-aided human translation* (MAHT) concerns machine aids for translators and the term Fully-*automated machine translation* (FAMT). Such as figure 2-1.

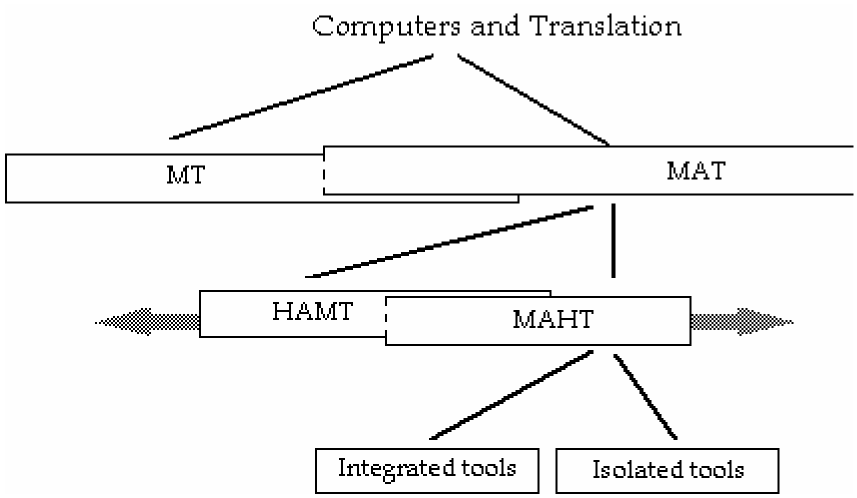


Figure 2-1 Computers and translation

Machine Assisted Translation

Machine Assisted Translation (MAT) is human translation supported by a computer system. Support is available by lexical data, grammatical help, translation memory, domain information and organizational support.

Fully-Automated Machine Translation

This is what most people have in mind when they think of using computers to translate. The idea is simple: feed in the text in one language and get the text out in another language. Unfortunately, the implementation of this idea presents obstacles that have
yet to be completely overcome.
The main problem is the complexity of language. Consider for instance the meanings of the word "can." Besides its use as a modal auxiliary verb,

"can" has several legitimate and slang meanings as a noun: container, depth charge, jail, and toilet. Not to
mention an archaic verb meaning "to know or to understand." Assuming that the foreign language has a separate term for each of these meanings, how is the computer to know which one to choose?

As it turns out, advances have been made in teaching computers to understand language based on context. More recent research is focusing on the use of probability theories for analyzing texts. But fully automated machine translation covering a broad range of subject areas is still a distant goal.

Human-Assisted Machine Translation

This form of computer-aided translation is currently a reality. The "human-assisted" part usually refers to editing texts before and after machine translation. Human translators make sure that a computer can understand source texts. After the machine translation, they edit the computer's rough (and sometimes humorous) output to provide a correct, idiomatic document in the target language. Alternately, some systems require a translator to work interactively with the computer as it is translating.
Human-assisted machine translation works best on texts employing a restricted vocabulary in a narrowly defined subject area. The Canadian government translates weather forecasts from English to French is one of the most successful implementations. The Pan American Health Organization relies heavily on machine translation for translating health-related texts from English to Spanish. Large computer companies use machine translation to varying degrees in translating their technical manuals.

The economy of human-assisted machine translation continues to be debated. The programs themselves are generally quite expensive; some require specialized hardware. Pre- and post-editing are skills that must be learned and that not all translators enjoy. Generating and maintaining databases of terms is time-consuming and may require special expertise. But for an organization that does large amounts of translation in a clearly defined subject area, human-assisted machine translation can be a cost-effective alternative to purely human translation.

Machine-Assisted Human Translation

This approach places the human translator at the center of the translation process, while attempting to provide him or her with tools that make the translation process more efficient and accurate. The tools available range from word processors to translation editors, comparison programs, and terminology management programs. Translation editors identify selected terms and suggest translations. Comparison programs speed the updating of existing translations by finding changes between versions of source documents. Terminology management programs allow translators to maintain flexible databases of terms in their subject area. Prices for such tools are reasonable enough for small translation departments and even individual translators to afford.

2.2 History of Machine Translation

The idea of using computers to translate or help translate human languages is almost as old as the computer itself (Trujillo A. 1999). Indeed, MT is one of the oldest non-numeric applications of computers. In history has been colorful and eventful, influenced by the politics, science and economic of different periods of modern history. It has been told many times, so only a brief summary follows.

Pre-computer: Some of the ideas that have influenced MT were already current or at least existent in the pre-computer era. Since at least the 17th century scholars and philosophers have proposed the use of language-neutral representations of meaning in order to overcome linguistic barriers. More recently, a mechanical procedure for carrying out translation was patented by the Russian Petr Smirnov-Troyanskii in 1933.

Initial efforts: Early proposal for use the numerical techniques in MT can be traced at least to 1947, when computers had just been successfully employed in deciphering encryption methods during the Second World War. A memo from Warren Weaver proposed specific strategies for using computers to translate natural languages. This memo initiated MT research in the USA and in the rest of the world, with the first public demonstration of a Russian-English prototype MT system in 1954. This event led to similar work in the then USSR and other places around the world.

The ALPAC Report (1966): The initial over-optimism in MT came to an end in the USA when the ALPAC report, commissioned by government sponsors of MT, suggested that MT was not cost-effective. The result was divergence of funding from MT and into AI and CL, with research continuing mainly outside the USA, although some groups there survived.

The 1970s and operational MT: Continued effort in MT yielded operational systems in the early 1970s. Systran began Russian-English translations for US Air Force in 1970, while Meteo began translating weather reports in 1976. Also in 1976 the commission of the European Union (then communities) installed an English-French version of Systran.

Rebirth in the early 1980s: The late 1970s and early1980s saw an increase in interest in MT. The Eurotra project from the European community began in the 1982, influenced by work done at Grenoble and Saabrucken since the 1960s and 1970s. Similarly, in Japan the Mu project started in 1982, and knowledge-based MT started in earnest in 1983 in the USA. Some commerical systems also began to appear.

Late 1980s and early 1990s: A number of companies, especially large Japanese electronics manufactures, began to market MT software for workstations. A number of products appeared for personal computers, and various MAT tools such as translation memory began to be commonly used. This period also saw the emergence of work on speech translation and of statistical approaches to machine translation.

One form that this takes is that of example-based machine translation (Furuse & Iida, 1992; Iida & Iida 1991; Nagao, 1992; Sato, 1992) in which a system of otherwise fairly conventional design is able to refer to a collection of existing translations. A much more radical approach, championed by IBM (Brown, Cocke, et al., 1990), is the one in which virtually the entire body of knowledge that the system uses is acquired automatically from statistical properties of a very large body of existing translation.

The most conspicuous example was the METEO system (Huchins and Somers, 1992), developed at the University of Montreal, which has long provided the French translations of the weather reports used by airlines, shipping companies, and others. Some manufacturers of machinery have found it possible to translate maintenance manuals used within their organizations (not by their customers) largely automatically by having the technical writers use only certain words and only in carefully prescribed ways.

Late 1990s and MAT: At the end of the decade we are seeing powerful translation engines on personal computers, translation on the Internet, widespread use of translation memory and translator’s workbenches, multimedia and software localization, as well as an increased interest in Example-based MT.

2.3 Approaches to Machine Translation

2.3.1 Traditional approaches

Traditionally approaches of MT can be classified by their architectures:

- Direct or transformer architecture
- Transfer based architecture
- Interlingual architecture

Figure 2-2 may be helpful in seeing the relation between interlingual, transfer and direct approaches.

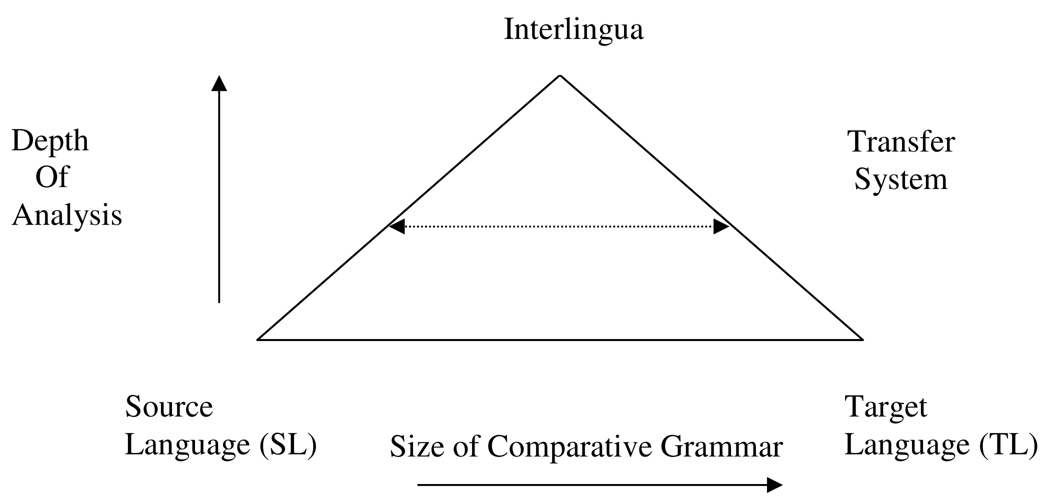


Figure 2-2: Traditional approaches of machine translation

A transfer component is a system of rules that relate words and structures in one language to words and structures of another. A transfer-based architecture or systems is simply one, which employs such components.

Interlingua systems dispense with such a component, mapping from the source language (SL) text to a more or less independent representation, and from that to the target language(TL) text. By comparison, the most abstract representation in a transfer based system are clearly either source, or target oriented, e.g. they may contain source or target lexical items. So called Direct systems resemble transfer systems in containing bilingual rules, but attempt to do without the abstract representation, mapping directly from SL to TL strings with only minimal structural analysis.

2.3.1.1 Direct or transformer architecture

The main idea behind transformed into output (target language) sentences by carrying out the simplest possible parse, replacing source words with their target language equivalents as specified in a bilingual dictionary, and then roughly re-arranging their order to suit the rules of the target language. The overall arrangement of such an Engine is shown in Figure 2-3

The first stage of processing involves the parser, which does some preliminary analysis of the source sentence. The result need not be a complete representation. This is passed to a package of rules which transform the sentence into a target sentence, using--where necessary--information provided by the parsing process. The transformation rules include bilingual dictionary rules and various rules to reorder words, for example, to make sure verbs have the correct person, number, and tense suffixes.

We can summarize the situation of the transformer engine architecture as follows:

- It is highly robust. That is, the Engine does not break down or stop in an 'error conditions' when it encounters input, which contains unknown words or unknown grammatical constructions. Robustness is clear important for general-purpose MT.
- In the worst case it can work rather badly, being prone to produce output that is simply unacceptable in the target language ('word salad’).
- The translation process involves many different rules interacting in many different ways. This makes transformer systems rather hard to understand in practice - which means that they can be hard to extend or modify.
- The transformer approach is really designed with translation in one direction, between one pair of languages in mind, it is not conductive to development of genuinely multi-lingual systems (as opposed to mere collections of independent one-pair, one direction engines).

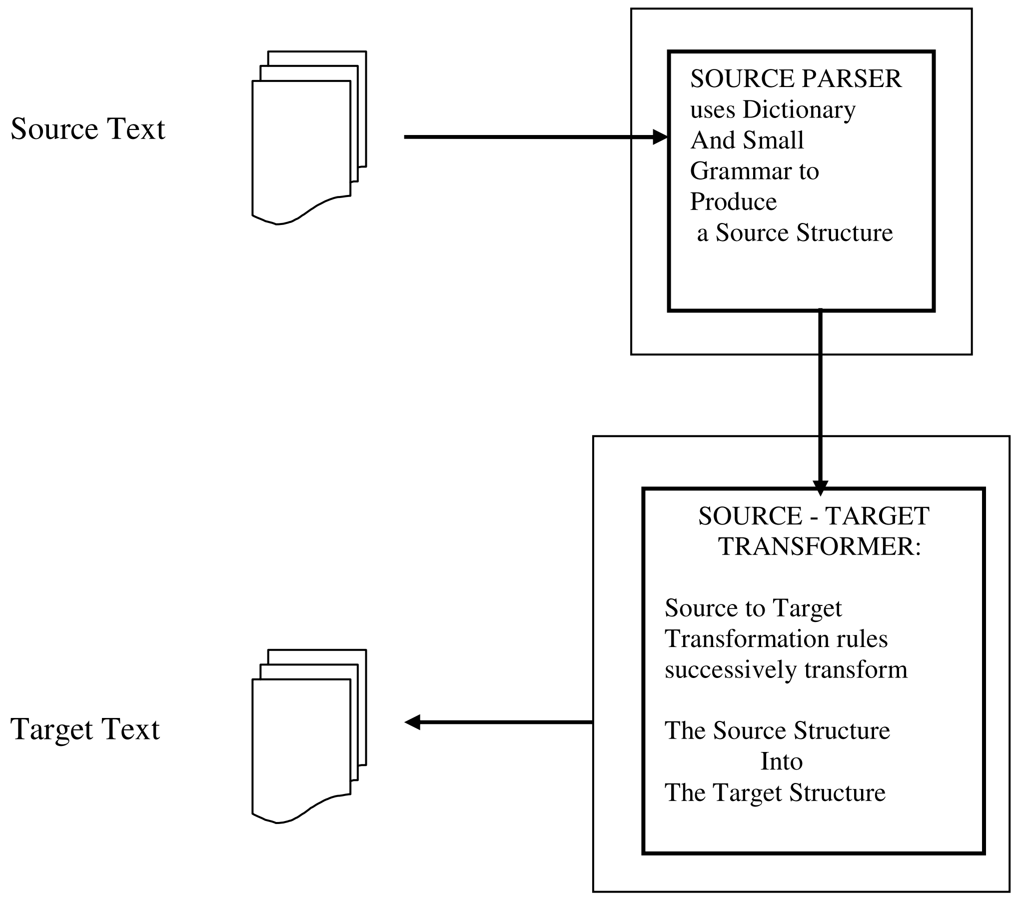


Figure 2-3: Transformer Architecture

2.3.1.2 Transfer-based architecture

The idea of transfer-based method is clear that if the system is translating from source language to target language, the first (analysis) step involves using the parser and the source grammar to analyze the input. The second (transfer) step involves changing the underlying representation of the source sentence into an underlying representation of the target sentence. The third (synthesis) step and final major step involves changing the underlying target representation into the target sentence, using a generator and the target grammar. The fact that a proper source grammar is being used means that the output of system-the target sentences are far more likely to be grammatically correct than those of transformer system. The components of a transfer system are shown in the following figure2-4.

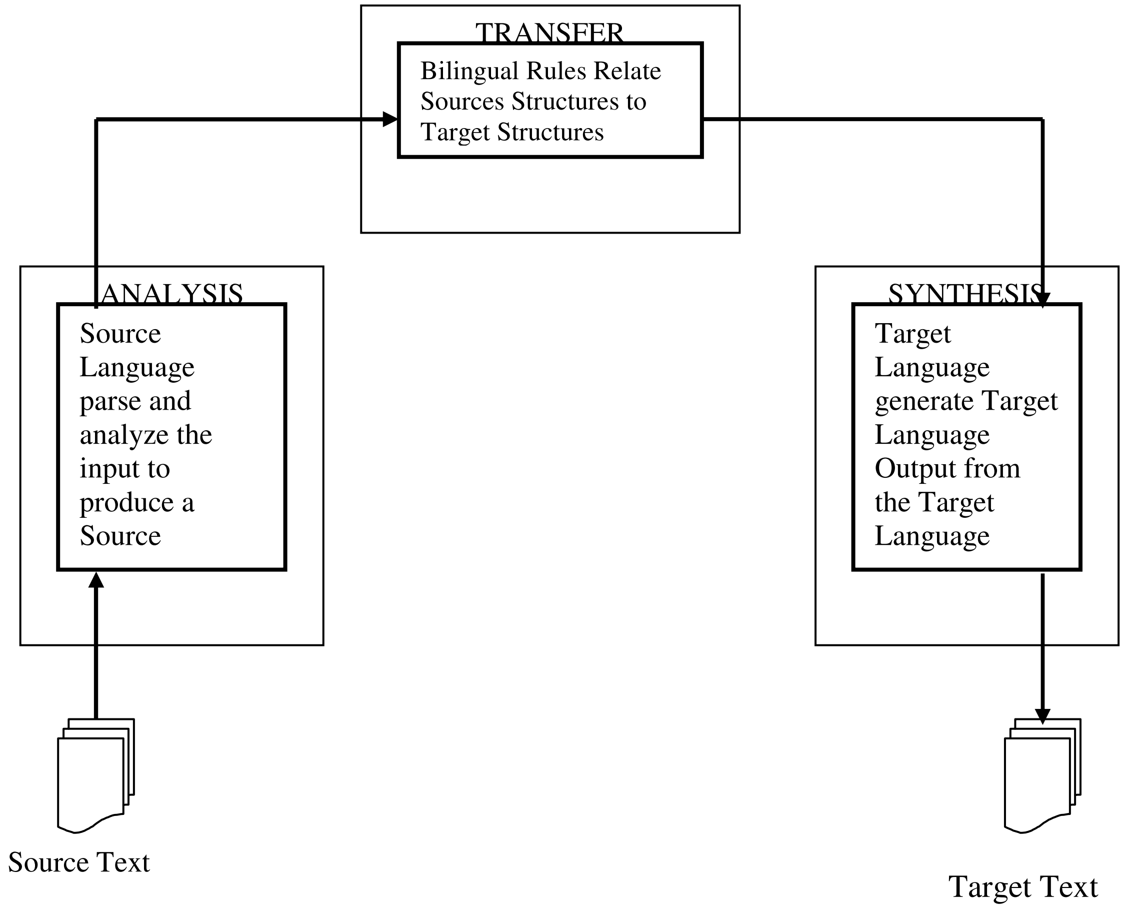


Figure 2-4: Components of a transfer system

Transfer architecture is the `standard model' for contemporary MT, and has been so for many years. Apart from a handful of commercially available systems which remain fundamentally `Direct', and a small number of Interlingual systems (notableyKBMT), the great majority of both research systems and development systems are transfer based. Some examples of Transfer systems are the following

(Several of these are excellently described in Hutchins and Somers, 1992): ARIANE (GETA) (Vauquois and Boitet, 1985) SUSY (Mass, 1987), METAL (Slocum, 1987, Bennett and Slocum, 1988), TAUM-AVIATION (Isabelle, 1987), ETAP-2(Apresjan et al, 1992), LMT (McCord, 1989), EUROTRA (Copeland et al, 1990),CAT-2(Sharp,1988), MIMO (Arnold and Sadler,1990), MIMO-2 (Van Noord et al, 1990), ELU (Estivalet al, 1990).

In a typical Transfer based system, the SL string is first analyzed morphologically,and then parsed to produce a surface syntactic representation. This is then transformed into a more abstract representation (RSL) so as to abstract away from phenomena with no translational relevance, and provide a more convenient representation of other information. Transfer converts this into a similarly abstract Target Language representation (RTL), from which synthesis produces a surface syntactic representation of the target text, and ultimately a TL string. The abstract level of representation is sometimes called an `Interface Structure' (IS), because it provides the interface between analysis and transfer, and transfer and synthesis. Deriving the abstract representation from the surface syntactic structure might involve abstracting away from differences in word order that have no semantic significance, or which are otherwise predictable

2.3.1.3 Interlingual architecture

The general idea suggested by the discussion of the pervious section is that the transfer system becomes simpler as linguistic analysis goes deeper - as the representations become more abstract. In fact, a major objective of MT research is to define a level of analysis which is so deep that the transfer system disappears completely. Given such a level of representation, the output of analysis could be direct input to the target synthesis component. Representation at such a level would have to capture whatever is common between sentences and their translations - that is they would have to be a representation of 'meaning '. Moreover, such a level of representation would have to be entirely language independent - for example, if it preserved features of the source language, one would still require a transfer component of some kind to produce the corresponding features of the target language. For this reason, such a level of representation is normally called an Interlingua and the system that uses such a level are called Interlingual. The components of a interlingual system are shown in the following figure2-5.

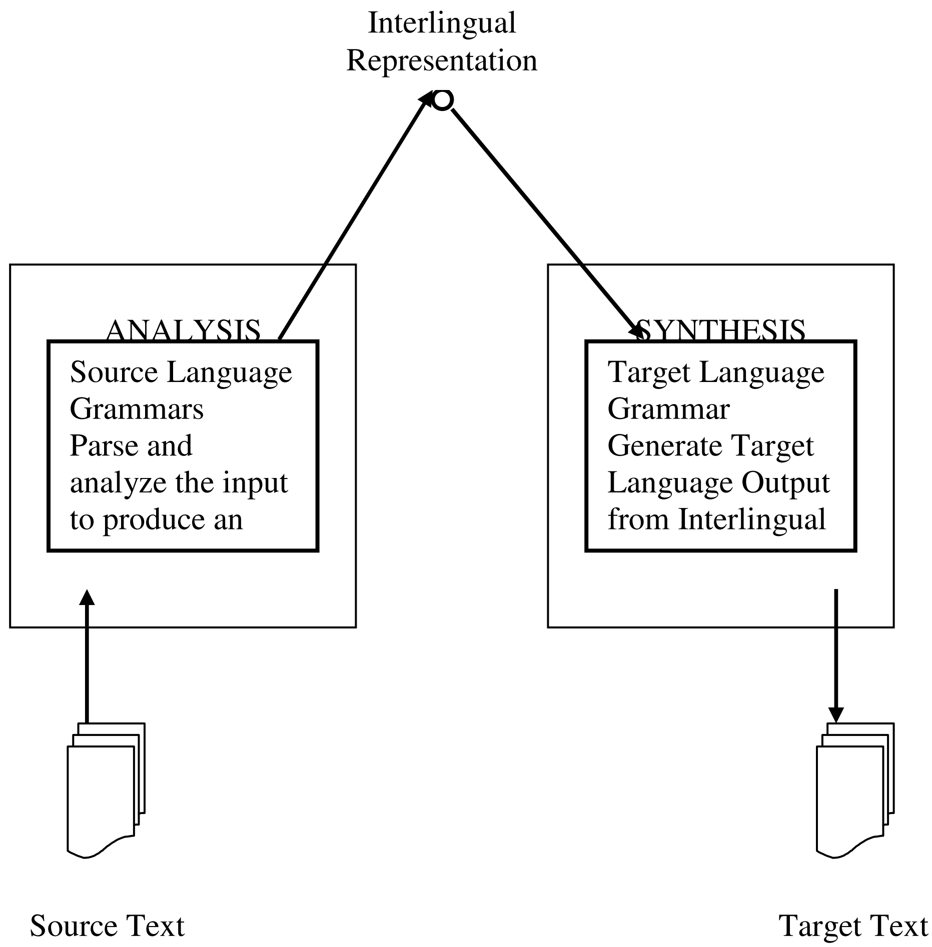


Figure 2-5 the Component of Interlingual System

There are a number of clear attractions to an interlingual architecture. First, from the scientific point of view, the idea of an interlingual is interesting and exciting. Second, from a more practical point of view, an interlingual systems promises to be much easier to extend by adding new language pairs, than a transfer system (or transformer system). This is because, providing the interlingua is properly designed, it should be possible to add a new language to a system simply by adding analysis and synthesis components for it. Compare this with a transfer system, where one needs not only analysis and synthesis, but also transfer components into all the other languages involved in the system.

The PIVOT system of NEC (Okumura, Muraki, et al., 1991; Muraki, 1989) and ATLAS II of Fujitsu (Uchida, 1989) are commercial systems among a number of research systems based on two-step method according to which texts are translated from the source language to an artificial interlingual representation and then into the target language. The Rosetta system at Phillips (Landsbergen, 1987), and the DLT system at BSO (Witkam, 1988; Schubert, 1988) in the Netherlands also adopt this approach.

2.3.2 Other approaches

A selection of approaches to MT that illustrate a range of useful techniques is presented. Other approaches of MT can be classified as follow:

- Knowledge-based approach
- Corpus-based approach

2.3.2.1 Knowledge-Based approach

The central principle underlying this approach is the stress on functionally complete understanding of the meaning of the source text as a prerequisite to successful translation. The term functionally(Nirenburg et al. 1992, p. 27) means that the meaning representation should be sufficient for translation to a number of languages, rather than sufficient for total understanding, which entails a more complete, human-like inferential process for understanding all implicit and explicit information.

Architecturally, knowledge-based machine translation systems belong to the class of interlingua-based systems, in which translation is basically a two-step process (analysis and generation). (This is in contrast to the class of systems that involve three steps -analysis, source-to-target language transfer and generation-and are therefore called transfer systems.) Example of interlingua systems that are not knowledge-based are CETA (Vauquois and Boitet 1985), DLT (Witkam 1983) and Rosetta (Landsbergen 1989). The main difference between such systems and knowledge-based machine translation ones is the expected depth of source language analysis and the reliance of KBMT systems on Explicit representation of world knowledge.

In the knowledge-based machine translation the first system was developed by Yorick Wilks at stanford University(Wilks 1973). Further experiments were conducted by Jaime Carbonell, Rich Cullingford and Anatole Gershman at Yale University (Cabonell et al. 1981) and Sergei Nirenbburg, Victor Raskin and Allen Tucker at Colgate University (Nirenburg et al 1985). Larger-scale development work followed, and a number of pilot knowledge-based machine systems have been implemented. The major efforts include ATALS-II (Uchida 1989 a and b), PIVOT (Muraki 1989), ULTRA (Farwell and Wilks 1991), the KBMT system for doctor patient communication (Tomita et al. 1987), KBMT-89 (Good man and Niernburg 1991) and Dionysus (e.g., Niernburg and Defrise in press, a and b; Meyer et al 1990; Calrson and Niernburg 1990). Some other systems (e.g.,HICATS/JE, Kaji 1989) are using some features of the knowledge-based approach.

Knowledge-based machine translation systems can be very profitably as research tools and testbeds in computational linguistics and artificial intelligence. To illustrate briefly:

- Knowledge-based machine translation provides an excellent tool for devising and testing new powerful specialized semantic-interpretation algorithms, such as noun- noun compound resolution or prepositional phrase attachment. With more types of semantic and pragmatic knowledge appearing in the knowledge representation, more specialized microtheories (Nirenburg and Defrise forthcoming, b; Nirenburg and Pustejovsky 1988) and the Mikrokosmos system of Onyshkevich and Nirenburg (1995) will be devised and/or incorporated into the process in order to deal with each phenomenon in a computationally tractable manner.
- An additional advantage of using knowledge-based machine translation as a research vehicle is that, being a comprehensive system, it allows immediate testing of new components, such as a new parser or a generator, in the context where actual output can be obtained and evaluated.
- The interface component of a KBMT system can serve as a medium for building other interfaces, notably for the purpose of computer-aided instruction and, in particular, for teaching foreign language. The interface can be also very useful in machine learning systems, especially those studying learning from text or learning by being told, or in systems that investigate hybrid learning processes that involve natural language.
- Components of a KBMT system can be individually recycled. The analysis module, for instance, can be useful for natural language interface to complex application such as expert system. More comprehensive understanding-and-generation

systems can also be used as components of a system modeling a cognitive agent- alongside other modules, such as planning and problem solving, perception and action -simulation components.

- Outside of machine translation proper, the technology developed for a knowledge-based machine translation system is readily usable in application that require different types of inputs and/or output to a natural language processor. Thus, instead of forwarding an interlingua text to the generator one can pass it on to a special reasoning program that will produce an abstract of the input text, or answer question based on it, or categorize the input text into one of a number of taxonomic classes (Hayes and Weinstein 1990). Knowledge-based machine translation systems can also be reconfigured for supporting natural language interfaces to database systems. Indeed, if a data manipulation (query) language is substituted for the inerlingua, the task of query formulation can become quite similar to that of analyzing a natural language input for translation.

2.3.2.2 Corpus-Based approach

Corpus-based machine translation is one of the alternative directions that have been proposed to overcome the difficulties of traditional systems. Two fundamental approaches in corpus-based MT have been followed. These are statistical and example-based machine translation (EBMT), also called memory-based machine translation (MBMT). Both approaches assume the existence of a bilingual text (an already translated corpus) to drive a translation for an input. While statistical MT techniques use statistical metrics to choose the most probable words in the target language, EBMT techniques employ pattern matching techniques to translate subparts of the given input (Arnold D., Balkan L., Humphreys R. Lee, Meijer S., Sadler L. 1994), (Trujillo A. 1999).

Corpus_based systems directly address the need for MT systems to be tuned to particular sublanguages or text types by using relevant previous translations. They also attemp to simplify knowledge capture by foregoing the need to manually develop some or all the linguistic resources needed for MT.

2.3.2.2.1 A Statistical approach

The idea of using statistics in MT dates back to the forties, but for obvious reasons - lack of appropriate computer power and of machine readable text corpora, etc. - it is just beginning to enter the field. Statistical analysis of natural language is faced with the number of problems of which we can address only a few here. Most notable is the issue of representativeness of text corpora (Sebba, 1991). Kucera et al. (1967) and Erk(1972) noticed that word frequency strongly depends on domain and/or genre of texts. This dependency is substantially stronger for noun than for verbs. Statistical methods have proven their value in automatic speech recognition (Bahl et al 1983) and have recently been applied to lexicography (Sinclair 1985) and natural language processing (Baker 1979); Ferguson 1980; Garside et al. 1987; Sampson 1986; Sharman et al. 1988).

Statistical MT systems (Trujillo A. 1999) rely on probabilistic and statistical models of the translation process trained on large amounts of bilingual corpora. Many of the models proposed include little or no explicit linguistic knowledge, relying instead on the distributional properties of words and phrases in order to establish their most likely translation. The general idea in statistical MT is that we look for features of bilingual corpus that are easily measured and see how these features can be used to predict translations. Features that can easily be measured include co-occurrence of two or more words in source and target text, relative position of words within sentences, length of sentences and many others; the idea is that these measures rely on little if any linguistic information.

One approach to statistical MT involves separate monolingual and bilingual sources of knowledge, which are combined to give the probability of a translation (Brown et al. 1990; Brown et al. 1993). In this work, there is a statistical language model that contains bilingual information. Translation then requires a method for:

(a) Computing the probability of a string being the translation of a SL string;

(b) Computing the probability of a TL string being a valid TL sentence,

(c) A technique to search for the TL string which maximizes these probabilities.

Mathematically, the relationship between these three processes may be expressed as:

$$\hat{t} = \underset{t}{\operatorname{argmax}}\, P(t)P(s|t)$$

The formula can be interpreted as saying that to translate source sentence **s,** we search for the target word string **t** that maximizes the value of the whole formula. The idea is that given sufficiently accurate statistics, the P(t) term biases the search towards grammatical TL word strings, while the P(s|t) term biases the search towards strings that are likely translations of the source sentence. This last conditional probability may appear confusing. Conditional is on the target word string, as it is easier to estimate the probability of a given source sentence from a TL word string than the other way round. To appreciate this, we can think of the source sentence as giving us hints about the TL sentence. Then it is simpler to estimate from corpora the probability of a set of hint (**s**) given a TL sentence (**t**) than it is to estimate the probability of a TL sentence from the hints alone.

The monolingual language model can be based on bigram or trigram models (Brown et al. 1992), from which the likelihood of a string of words being a valid sentence can be computed. By contrast, the translation model uses the frequency of co-occurrence of source and target words, the length of the sentences in which they appear, their positions within their respective sentences, the fertility of the TL word (the number of SL words from which it arises), the actual words from which a TL word derives, and the position of these SL words in the SL string. (Brown et al. 1993) propose a series if increasingly more sophisticated models that include more and more of these features.

2.3.2.2.2 Example-Based approach

Example-based machine translation (EBMT) is based on idea of performing translation by imitating translation examples of similar sentences (Nagao 1984). In this type of translation system, a large amount of bi/multi-lingual translation examples has been stored in a textual database and input expressions are rendered in the target language by retrieving from the database that example which is most similar to the input. There are three key issues, which pertain to example-based translation:

- establishment of correspondence between units in a bi/multi-lingual text at sentence, phrase or word level
- a mechanism for retrieving from database the unit that best matches the input
- exploit the retrieved translation example to produce the actual translation of the input sentence

(Brown 1991) and (Gale 1991) have proposed methods for establishing corresponding between sentences in bilingual corpora. (Brown 1993), (Sadler 1990) and (Kaji 1992) have tackled the problem of establishing correspondences between words and phrases in bilingual texts. Third key issue of EBMT, that is exploiting the retrieved translation example , is usually dealt with by integrating into the system conventional MT techniques(Kaji 1992), (Sumita 1991). Simple modification of the translation proposal, such as word substitution, would also be possible, provided that alignment of the translation archive at word level was available.

Exemplar-based representation has been widely used in machine learning (ML). According to(Medin and Schaffer 1978) who originally proposed exemplar-based learning, example are stored in memory without any change in the representation. Here, an exemplar-based learning is to use past experiences or cases to understand, plan, or learn from novel situations (Hammond, K.J. 1989), (Kolodner, J.L. 1988) and (Ram, A. 1993).

EBMT has been proposed by (Nagao, M. A 1985) as translation by Analogy, which is in parallel with memory, based reasoning, case-based reasoning and derivational analogy. Example-based translation relies on the use of past translation examples to derive a translation for a given input. The input sentence to be translated is compared with example translations analogically to retrieve the closest examples to the input. Then, the fragments of the retrieved examples are translated and recombined in the target language. Prior to the translation of an input sentence, the correspondences between the source and target languages should be available to the system ; however this issue has not been given enough consideration by the current EBMT systems. (Kitano, H. 1993) has adopted the manual encoding of the translation rules, however this is a difficult and an error-prone task for a large corpus. (Wu, D. 1995) uses a method to extract phrasal translation examples in sentence-aligned parallel corpora using a probabilistic translation lexicon for languages pair.

2.4 Conclusion

The introduction in the 1960s of `Transfer based' approaches with significant amounts of linguistic analysis was widely regarded as signaling a `Second Generation' of systems (following the First Generation `Direct' systems). Since then, there has been speculation as to what the `Third Generation' would be, and when it would arrive. However, despite the continued operational success of descendents of `Direct' systems, and significant progress in Interlingual approaches, the advantages of a Transfer architecture (in particular, feasibility, and the ability to exploit the most up-to-date computational linguistic techniques) mean that it is likely to remain the `standard model' of MT for the forseeable future. However, this is not in itself a particularly interesting, since, as we have observed, the architecture is general enough to encompass a number of significantly different views, and to permit a wide range of techniques to be applied.

There are many factors, which make transfer an attractive design for MT(Trujillo A. 1999).

- Many systems are bilingual, or their principal use is for translation in one direction between a limited number of languages.
- Where full multilinguality is required, it is possible to have a hub language into and out of which all translation is done.
- Portions of transfer modules can be shared when closely related languages are involved. For example, an English-Portuguese

module may share several transformations with an English-Spanish module.

- Simpler analysis, faster development of grammar, simpler automatic grammar induction.

As shown, transfer is still a successful approach sufficient enough to develop MT systems. This enables us to design and implement a MT system for translating English noun phrases to Arabic.

CHAPTER 3

A TRANSFER-BASED MT APPROACH FOR ENGLISH NOUN PHRASES

This chapter proposes a transfer-based approach that handles the translation of English noun phrases into Arabic. It includes three main modules, responsible for analysis, transfer, and generation. Section 3.1 outlines the overall architecture of the proposed English to Arabic MT system. In section 3.2 we describe how the analysis component assigns grammatical structures to the input noun phrase by means of English grammatical rules and an English monolingual dictionary. In Section 3.3 we describe how the transfer component builds target language equivalents of the source language grammatical structures by means of a comparative grammar that relates every source language representation to some corresponding target language representation. In Section 3.4 we describe how the generation component provides the target language translation.

3.1 Overall Structure of the System

The architecture of the transfer-based English to Arabic MT system is given in Fig. 3.1. In this figure, the arrows indicate the flow of information. Oval blocks represent the basic modules of the system.

Rectangle blocks represent the linguistic knowledge. The proposed system is based on the transfer approach (Arturo Trujillo 1995), (Turcato, Laurens, McFetridge, and Popowich 1997), (Volk and Oberholzer 1996), with three main components: an analyser, a transfer component, and a generation component. The following box summarises the translation process.

1. **Pre-processor will read in a noun phrase, returning the words in the noun phrase to a list of words.**
2. **The morphological analyzer provides information about inflectional features as well as the primitive-form of an inflected word.**
3. **The syntactic parser builds a syntactic dependency tree, which represents meaningful linguistic relationships between constituents of a noun phrase.**
4. **Lexical transfer will map English lexical units to their Arabic equivalent. It will also map English morphological features to the corresponding set of Arabic features.**
5. **Structure transfer will map the English dependency tree to the equivalent Arabic syntactic structure.**
6. **The Arabic generator will synthesis the inflected Arabic word-form based on the morphological features and traverse the syntactic tree to produce the surface Arabic noun phrase.**

The system is implemented in Prolog and the parser is written in DCG formalism. DCG translate grammar rules directly to Prolog, producing a simple top-down parser. So all it is needed is a way to reformulate grammar rules as clauses in DCG.

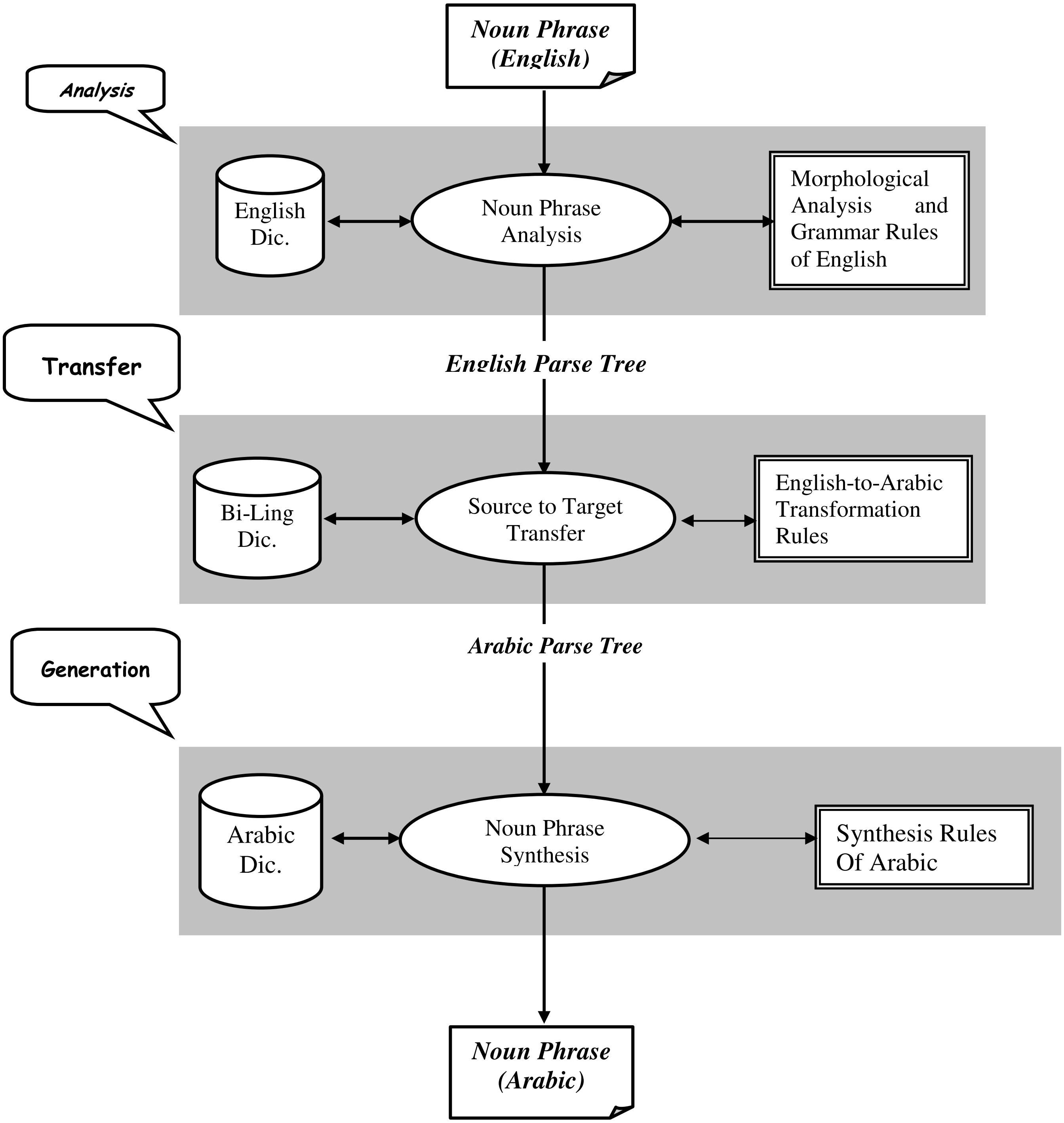


Fig 3-1 Overall Structure of English-Arabic Noun Phrase

3.2 The Analysis Components

English noun phrases are frequently used in scientific and technical documents. In a recent study conducted by Quinn (1997), he mentioned that

> *"Justeson and Katz (1995) analysed around 200 English technical terms from each of four domains. They found that about 96% of the terms were noun phrases (NPs), about 4% adjectives, and less than 0.5% (3) were verbs. The NPs had a mean length of 1.9 words - 70% were compound. 97% of the compound NPs contained only nouns and adjectives, and 3% included a preposition (almost always "of"). A terminology identification algorithm was presented, which includes the recognition of the standard patterns for terms. A candidate term is either 'a string of nouns and/or adjectives, ending in a noun', or 'two such strings, separated by a single preposition'. This gives seven basic patterns for two or three word terms, containing only nouns (N), adjectives (Adj) and prepositions (Prep). 95% of the compound NPs analysed were of length two or three."*

These results have motivated us to choose English noun phrase as the source text for the translation system. We found that the translation of the title of scientific texts, e.g. theses and journals, to be closely applicable to this system.

The analysis step comprises English dictionary, English morphological analyzer, and English noun phrase parsing.

3.2.1 Linguistic Background

The parsing of English noun phrase produces the linguistic information about the sentence structure and each word in the sentence (Bond, Ogura, and Kawaoka 1996). So, we will begin our discussion by presenting the aspects of the sentences and words in English. Then, we focus our attention to the design of the parser.

The Sentence Structure

The structure of a sentence indicates the way that words in the sentence are related to each other. This structure indicates how the words are grouped together into phrases, what words modify what other words, and what words are of central importance in the sentence. In addition, this structure may identify the types of relationships that exist between phrases and can store other information about the particular sentence structure that may be needed for later processing. This information is often presented in a tree form, such as shown in Fig. 3.2, which shows the structure for the English sentence "Rice flies like sand." The sentence is formed from a noun phrase describing a type of substance, "rice", and a verb phrase stating that this substance flies like sand (say, if you throw it). The structures also gives further details on the structure of the noun phrase and verb phrase, and identify the part of speech for each word. In particular, the word "like" is a preposition.

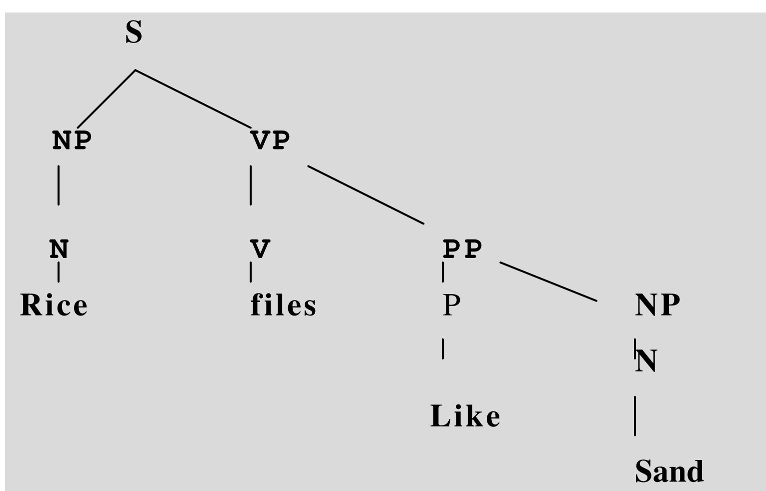


Fig.3-2 The structure of English sentence

The Nature of Words

Words fall into two major groups: the *content words*, which serve to identify objects, relationships, properties, actions and events in the world, and the *function words*, which serve a more structural role in putting words together to form sentences. There are four general classes of content words, based on what the words tend to describe.

- **Nouns** describe classes of objects, events, substances, and so on (for example, ball, man, sand, and idea)
- **Adjectives** describe properties of objects (for example red, tall, and special)
- **Verbs** describe relationships between objects, activities, and occurrences (for example seems, ate, and laughing).

  Note that:
- –ing form of a verb, called *gerund*, can be used as a noun (for example modeling, programming, and engineering).
- –ed form of a verb, i.e. past participle form, can be used as an adjective (for example integrated, isolated, and distributed).
- **Adverbs** describe properties of relationships or other properties (for example very and slowly)

The function words tend to define how the content words are to be used in the sentence, and how they relate to each other. Some common classes of function words are described in the following.

- **Determiners** indicate that a specific object is being identified (for example, a, the, an, this, and that)
- **Quantifier** indicate how many of a set of object are being identified (for example, all, many, some, and none)
- **Prepositions** signal a specific relationship between phrases (for example, in, onto, by, and through)
- **Connectives** indicate relationships between sentences and phrases (for example, but, and while).

English often allows a word in one class to be converted into a word in another class fairly freely by adding suffixes or by its particular use in a sentence. For instance, the word *sugar* is a noun, but like many nouns it can be used as a verb (for example, *He sugared the coffee*) or as an adjective (for example, *It was too sugar for me*). This is why morphology is so important in any theory of language. We will assume throughout, however, that all the uses of a particular word will be defined in advance in a structure called the *dictionary*.

3.2.2 The English Dictionary

Some authors use the word lexicon instead of dictionary. Lexicon is a technical term used to refer to call *morphemes* ("root forms"). A morpheme is the minimal meaningful unit in a language. Some authors use lexicon to refer to the vocabulary of a language as used by an individual speaker.

Dictionaries are the largest components of a MT system in terms of the amount of information they hold. If they are more than simple word lists (and they should be, if a system is to perform well), then they may well be the most expensive components to construct. More than any other component, the size and quality of the dictionary limits the scope and coverage of a system, and the quality of translation that can be expected.
The dictionaries are where the end user can expect to be able to contribute most to a system—in fact, an end user can expect to have to make some additions to system dictionaries to make a system really useful. While MT supplies rarely make it possible for users to modify other components, they normally expect them to make additions to the dictionary. Thus, from the point view of a user, a basic understanding of dictionary construction

and sensitivity to the issues involved in 'describing words' is an important asset.

Normal dictionaries, are collections of entries such as "these". That is, they are basically lists of words, with information about the various properties. While grammar rules define all the possible linguistic structures in a language, the descriptions of individual words that are found in the dictionary or dictionaries state which words can appear in which of the different structures. A common (though not completely correct) view is that dictionaries contain all the 'irregular', or unpredictable information about words, while grammars provide general rules about classes of word, and phrases (this is only true if one exclude morphological rules and idioms from the dictionary–the former can be viewed as dealing with classes of words, and the latter are phrases).
The following describes the structure of entries for the proposed monolingual English dictionary. Table 3-1 shows examples of these entries.

- *Noun*: a content word which has two features the stem-form and the number that takes either a singular or plural. Nouns are stored in singular form. Irregular plurals have additional entries.
- *Adjective*: a content word, which has a stem-form feature.
- *Quantifier*: a function word, which has a stem-form feature.
- *Separator*: a function word, which has a stem-form feature. This includes connectives, prepositions, and special words that are used as a separator of a compound noun phrase.

Table 3-1 Examples of the monolingual English dictionary entries

Noun		**Adjective**	**Quantifier**	**Separator**
Form	***Number***	***Form***	***Form***	***Form***
Knowledge	Singular	neural	Some	and
Base	Singular	relational	On	of
Querying	Singular	integrated	Towards	":"

The lexicon must contain information about all the different words that can be used. When a word is ambiguous, it may be described by multiple entries in the lexicon, one for each different use (Melamed 1998), (Turcate 1998) and (Rayner, Carter, Bretan, Eklund, Wiren, Leo Hansen, Andersen, Philp, Finn, and Thomsen 1997).

The dictionary is implemented as Prolog facts. Figure 3.3 shows the implementation of the examples given in Table 3-1.

noun(knowledge,sg).	**adj(neural).**	**Quantifier(some).**	**separator(and).**
noun(base,sg).	**adj(relational).**	**Quantifier(on).**	**separator(of).**
noun(querying,sg).	**adj(integrated).**	**Quantifier(any).**	**separator(':').**

Fig.3-3 Implementation of the English dictionary

3.2.3 The English Morphological Analyzer

Morphological analysis (Gasser 1995) is the identification of a stem-form from a full (inflected) word-form (and sometimes also the identification of the syntactic category of the stem). A morphological analyzer must be able to undo the spelling rules for adding affixes. For example, the analyzer must be able to interpret "moved" as "move" plus "ed". For English, a few rules cover the generation of plurals and other inflections such as verb endings. Since verbs are irrelevant to noun phrases, we ignore the treatment of verbs in our discussion from now on.

Analysis of Plurals

The following words obey regular rules for the generation of plurals.

```
problems → problem + s
source → source + s
approaches → approach + es
communities → community + ies
```

Most English nouns, for example, use the same set of suffixes to indicate different forms: -s added for the singular, -es for the nouns are ending with (sh, ch, s, z, x, o) and so on. Without any morphological analysis, the lexicon would have to contain every one of these forms.

The remaining words are irregular.

```
mice → mouse
fish → fish
rooves → roof + ves
```

These forms are being explicitly included in the lexicon.

The only morphological analysis that can be done here is to recognize a single noun form from its plural. The architecture of the morphological analyzer is given in Figure 3.4. The implementation of the morphological analyzer is given in Fig. 3.5. This figure shows that the morphological analysis proceeds as follows:

1. Lookup the stem in the dictionary, if it is found return the stem and its features.
2. Otherwise, analyze the word to get its singular form:
   - Examine the ending of a word,
   - Compare it with the standard endings,
   - Derive the noun stem that could be consistent,
   - Compare this stem with a dictionary of stems, and
   - Return its number feature as plural.

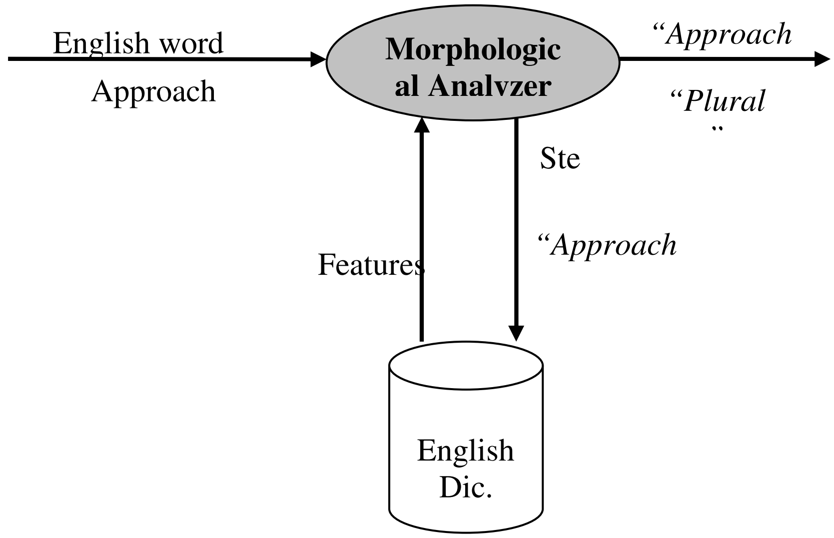


Fig. 3-4 Architecture of the morphological analyzer

```
morph(EWord,EWord,Numero):-
        noun(EWord,Numero),!.
morph(EWord,Single,Numero):-
        name(EWord,W1),
        pl_to_sg(W1,W2),
        name(Single,W2),
        noun(Single,sg),!,
        Numero=pl.
pl_to_sg(EWord,Single):-
        pl_to_sg_suffix(PLEnding,SGEnding),
        append(S,PLEnding,EWord),
       append(S,SGEnding,Single).
% PLURAL-ARG1+ARG2 → stem
pl_to_sg_suffix("s","").
pl_to_sg_suffix("es",""). % for endings sh, ch, s, x, z, o
pl_to_sg_suffix("ies","y").
pl_to_sg_suffix("ves","fe").
pl_to_sg_suffix("ves","f").
pl_to_sg_suffix("es","is").
pl_to_sg_suffix("a","um").
pl_to_sg_suffix("on","a").
pl_to_sg_suffix("i","us").
pl_to_sg_suffix("ae","a").
pl_to_sg_suffix("ices",ex).
pl_to_sg_suffix("ices",ix).
```

Fig. 3-5 Implementation of the morphological analyzer of Nouns

3.2.4 The English Noun Phrase parser

The development of the parser is a two-step process. In the first step, we acquire the rules that constitute a grammar for the English noun phrase that gives a precise account of what it is for a noun phrase to be grammatically correct. The grammar covers the titles of theses and journals from the domain of computer science. The second step is to implement the parser that assigns grammatical structure onto input noun phrase.

The Grammar

The grammar of noun phrases is acquired from the analysis of 50 thesis titles (See Appendix A). Table 3-2 shows some examples (up to 4 words long) of titles that illustrate the basic patterns that constitute the grammar of a *simple noun phrase*.

Table 3-2 Pattern of a simple noun phrase

Example	Pattern
Scheduling	N
software maintenance	N N
Neural Networks	Adj N
Knowledge base verification	N N N
natural language interface	Adj N N
Integrating hierarchical classification	N Adj N
Feedforward neural networks	Adj Adj N
Arabic syntax analysis environment	N N N N
graphical user interface application	Adj N N N
Integrating modern information technology	N Adj N N
Power transformer differential relay	N N Adj N
Picture based hierarchical classification	N Adj Adj N
Concurrent abductive logic programming	Adj Adj Adj N

From Table 3-2, *Grammar 1* that generates simple noun phrases could be written as follows:

(1) NP → N
(2) NP → N NP
(3) NP → Adj NP ***(Grammar 1)***

It is common to use determiners and quantifiers with noun phrases, see examples given in Table 3-3.

Table 3-3 Pattern of a simple noun phrase with a determiner and quantifier

Example	**Pattern**
An Arabic syntax analysis environment	Det NP
A natural language interface	Det NP
A knowledge representation scheme	Det NP
The knowledge level architecture	Det NP
The irrigated wheat testbed	Det NP
On the recent approaches	Quantifier NP
Some computer aided software engineering tools	Quantifier NP
From dependency networks to KADS	Quantifier NP
On integrated agricultural expert systems	Quantifier NP
Towards a knowledge representation language	Quantifier Det NP
On the recent approaches	Quantifier Det NP

From Table 3-3, *Grammar 2* could be drawn form *Grammar 1* by adding rules (4-5) that allows for using determiners and quantifiers with noun phrases.

(1) NP → N
(2) NP → N NP
(3) NP → Adj NP
(4) NP → Det NP
(5) NP → Quantifier NP

(Grammar 2)

Our analysis indicate that noun phrases either occurs in a simple form or in a compound form. *Grammar 2* describes the simple form of a noun phrase. The compound form of a noun phrase is two or more noun phrases separated by a connector, a preposition, or a special word, collectively called separators. Special words (SW) are either the special symbol (":"), colon, or a word that can be used to recognize the beginning of a new noun phrase. Table 3-4 shows examples of a compound noun phrase.

Table 3-4 Pattern of a compound noun phrase

Example	Pattern
Integrating the hierarchical classification **and** intelligent data base generic tasks	NP connector NP
Natural resources conservation **and** crop management expert systems	NP connector NP
Problems **in** software maintenance	NP Prep NP
An expert system **for** seedling production management	NP Prep NP
Towards a knowledge representation language **based on** open architecture model	NP SW NP
Use **of** expert system **in** irrigation practices	NP Prep NP Prep NP
Assessment **of** structure **using** neural networks approach	NP Prep NP SW NP
Agricultural expert systems **:** development **and** practice	NP SW NP connector NP
Comprehensive study **on** crisis management **with** an application **on** real problem	NP Prep NP Prep NP Prep NP
Design **and** implementation **of** a system **for** program verification	NP connector NP Prep NP Prep NP

From Table 3-4, *Grammar 3* that describes compound noun phrases could be drawn form *Grammar 2* by adding the rule 6 and 7.

(1) NP → N
(2) NP → N NP
(3) NP → Adj NP
(4) NP → Det NP
(5) NP → Quantifier NP
(6) NP1 → NP Sept NP1
(7) NP1→ NP ***(Grammar 3)***

The Parser

The task of the noun phrase parser is to determine the structure of a noun phrase and to construct a parse tree that represents this structure. The architecture of the noun phrase parser is shown in Fig. 3-6. In this architecture, depending on the syntactic category, the word is recognized either by calling the morphological analyzer to identify a noun or directly by looking up the word in the dictionary.

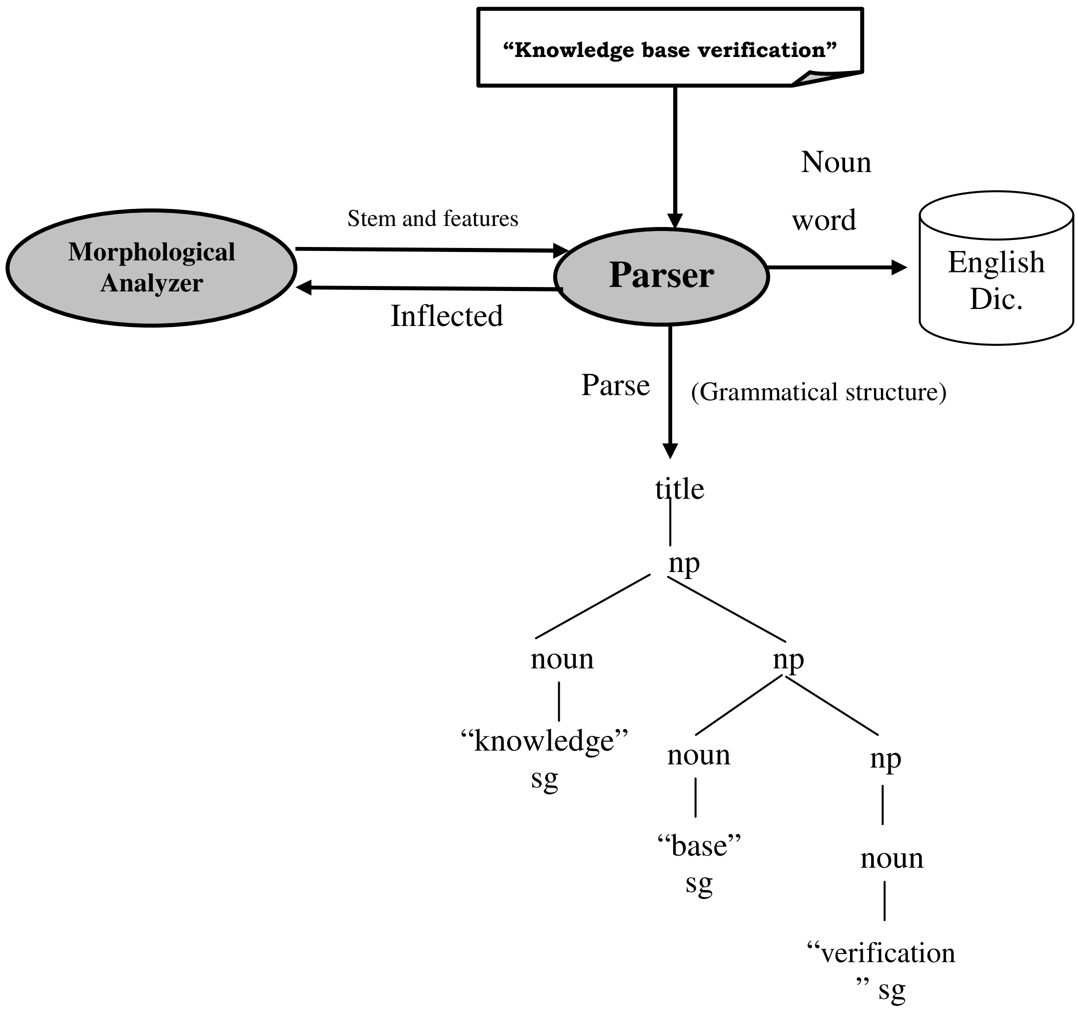
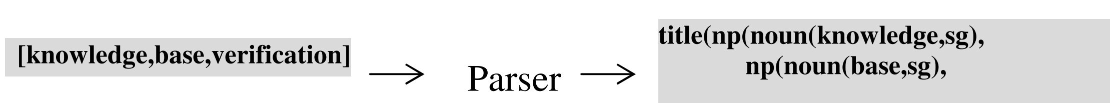


Fig. 3-6 Architecture of the noun phrase parser

Many Prolog implementations (Ball, 1995) provide a notational extension called DCG. This makes it very easy to implement formal grammars in Prolog. A grammar stated in DCG is directly executable by Prolog as a syntax analyzer. Figure 3-7 shows the implementation of *Grammar 3* in DCG. The function of the parser is to take as its input a list of words, represented as a Prolog list, which constitutes a noun phrase and produces as its output the parse tree, represented as a Prolog term.



```
title(title(NP))            -->   np1(NP).
np1(np1(NP,Sept,NP1))       -->   np(NP),sept(Sept), np1(NP1).
np1(NP)                     -->   np(NP)

np(np(Det,NP))              -->   determiner(Det), np(NP).
np(np(ADJ,NP))              -->   Adjective(ADJ), np(NP).
np(np(Q,NP))                -->   quantifier(Q), np(NP).
np(np(Noun,NP))             -->   noun(Noun), np(NP).
np(np(Noun))                -->   noun(Noun).
noun(Noun)                  -->   [Word],   {morph(Word,Stem,Numero),!,
                                            Noun=noun(Stem,Numero)}.
adjective(Adj)              -->   [Word],   {adj(Word),!, Adj=adj(Word)}.
sept(Sept)                  -->   [Word,P], {separator((Word,P)),!,
                                            Sept=sept((Word,P))}.
sept(Sept)                  -->   [Word],   {separator(Word),!,
                                            Sept=sept(Word)}.
quantifier(Q)               -->   [Word],   {quantifier(Word),!,
                                            Q=quantifier(Word)}.
determiner(det(a))          -->   [a].
determiner(det(an))         -->   [an].
determiner(det(the))        -->   [the].
```

Fig. 3-7 Implementation of the noun phrase parser

Recall that our MT approach is based on a tree-to-tree transformation. So, it is useful to have the parse tree explicitly represented in the program. The parse tree can be easily constructed using arguments of non-terminals in the DCG notation. We can conveniently represent a parse tree by a Prolog term whose functor is the root of the tree and whose arguments are the subtrees of the tree. For example, the parse tree for the noun phrase "knowledge base" would be represented as:

np(noun(knowledge,sg),np(noun(base,sg)))

A useful notational extension provided by DCG that allows us to insert normal Prolog goals into grammar rules. Such goals have to be enclosed in curly brackets to make them distinguishable from other symbols of the grammar. This is used to implement the interaction between the morphological analyzer and the parser as follows:

noun(Noun) --> [Word],

{morph(Word,Stem,Numero),!,Noun=noun(Stem,Numero)}.

In order to make our system convenient, we have implemented a pre-processor that reads in a noun phrase, returning the words in the noun phrase to a Prolog list. This is shown in Fig. 3.8.

```
% Reads in a sentence, binding the words in the sentence to the list
Words.
read_sent(Words) :-
      get0(Char),
      read_sent(Char,Words).
read_sent(C,[]) :- newline(C),!.
read_sent(C,end_of_file) :- eof(C),!.
read_sent(C,Words) :-
      skipchar(C),!,
      get0(Char),
      read_sent(Char,Words).
read_sent(Char,[Word|Words]) :-
      read_word(Char,Chars,Next),
      name(Word,Chars),
      read_sent(Next,Words).
skipchar(13).       % skip linefeed char
skipchar(32).       % skip spaces
newline(10).
eof(-1).
read_word(C,[],C) :- skipchar(C),!.
read_word(C,[],C) :- newline(C),!.
read_word(Char,[Char|Chars],New) :-
      get0(Next),
      read_word(Next,Chars,New).
```

Fig.3-8 Implementation of the Pre-Processor of the noun phrase parser

3.3 The Transfer Components

Broadly speaking, a *transfer component* is a system of rules that relate words and structures in one language to words and structures of another language (Arnold, 1994). In our noun phrase translator, the actual translation occurs in the transfer phase: a second set of rules that tells how to convert the English parse tree into an Arabic syntactic structure, represented as a Prolog term.

```
title(np(noun(knowledge,sg),
      np(noun(base,sg),
         np(noun(verification,sg)))))   ⟶ Transfer ⟶   title(np(noun(تحقيق,sg),
                                                              np(noun(قاعدة,sg),
                                                                 np(noun(معرفه,sg)))))
```

The English to Arabic transfer involves two steps:

1. *Lexical transfer*. This maps English lexical units to their Arabic equivalent. It also maps English morphological features to the corresponding set of Arabic features.
2. *Structural transfer*. This maps the English parse tree to the equivalent Arabic syntactic structure.

This transfer process is carried out with the help of a bi-lingual dictionary and a termbase, which is a gerund terminology.

3.3.1 English to Arabic Transfer: The Comparative Grammar

We can recognize the syntactic system of a natural language as a set of rules applied in the composition of sentences and clauses. This is what so-called grammar. Grammatical studies rest, essentially, upon two bases. First, lexical units (morphemes) are enumerated and classified. Second, interrelationships between morphemes should be classified and studied. In transfer approach, we try to relate the syntactic system of English with the syntactic system of Arabic by means of comparative grammar that maps lexical units and syntactic structures of English to that of Arabic.

Lexical Transfer

Just as each monolingual grammar of both source language and target language has a dictionary of rules (e.g. N → knowledge). These grammars are used by parsers to analyze sentences in each language into representations, which show their underlying structure, and by generators to produce output sentences from such representation.

The comparative grammar also has bilingual dictionary rules. These relate every source sentence representation to some corresponding target language representation –a representation which will form the basis for generating a target language translation:

Knowledge ⇔ معرفه
Base ⇔ قاعدة
Verification ⇔ تحقيق

The ⇔ in the present rules indicate that they can (in principle) serve as English-Arabic or Arabic-English rules.

The features on the node of the representation must also be translated in some way. Features are pairs that consist of an attribute, such as *number*, *definition*, and a value such as *singular.* The rules relevant to noun phrase are straightforward, indicating that the given values are simply carried over from source structure to target structure:

{number = sg} ⇔ {number = sg}

These dictionary rules can be seen as relating leaves (the word nodes) on the English parse tree to leaves on the target Arabic tree.

title(np(noun(knowledge,sg),
np(noun(base,sg), → Lexical Transfer → **title(np(noun(معرفه,sg),**
np(noun(قاعدة,sg),

Structural transfer

The comparative grammar also contains some structural rules, which relate other parts and nodes of the two trees to each other.

There is a relationship between adjacent units in the English noun phrase. This concerns the preserved order, the relative positioning (precedence), of lexical units in the noun phrase. For example, adjective in English precedes its noun, like in (good man), while in Arabic, noun precedes the adjective. Consequently, restructuring of the English parse tree is needed to conform to the target Arabic grammar. The transfer rules described here deal with the restructuring of the parse tree and reordering of words.

title(np(noun(معرفه,sg), **np(noun(قاعدة,sg),**	→ Structural Transfer →	**title(np(noun(تحقيق,sg),** **np(noun(قاعدة,sg),**

In addition to the lexical transfer rules that relate leaves on the English parse tree to leaves on the target Arabic tree, the comparative grammar also contains some structural rules which relate other parts and nodes of the two trees to each other.

In the structural rules that follows, the LHS describes an English structure and the RHS describes the Arabic, and \$1, \$2, …\$k are variables interpreting as standing for pieces of English structure on one side, and for their translations on the other side. Arnold (1994) introduced the method of defining the structure transfer rules described here.

The main difference between the English and Arabic parse tree representation is that words are in reverse order. A relatively simple straightforward example where a more complex example is called for involves the translation of "networks performance evaluation" into "تقييم أداء شبكة", which shows the switching of words. Such a rule might looks as follows:

(1)	$[w_i{:}\$1,\ w_{i+1}{:}\$2,\ \ldots,\ w_k{:}\$k]$	$(1 \le i \le k)$
	$\Leftrightarrow$	
	$[w_k{:}\$k,\ w_{k-1}{:}\$k\text{-}1,\ \ldots,\ w_i{:}\$i]$	$(1 \le i \le k)$

This rule says that the translation of the word at level i is switched with the word at level k-i+1. Where k is the number of noun phrases equivalent to maximum (sub)tree level. This process of word switching is illustrated in Fig. 3.9.

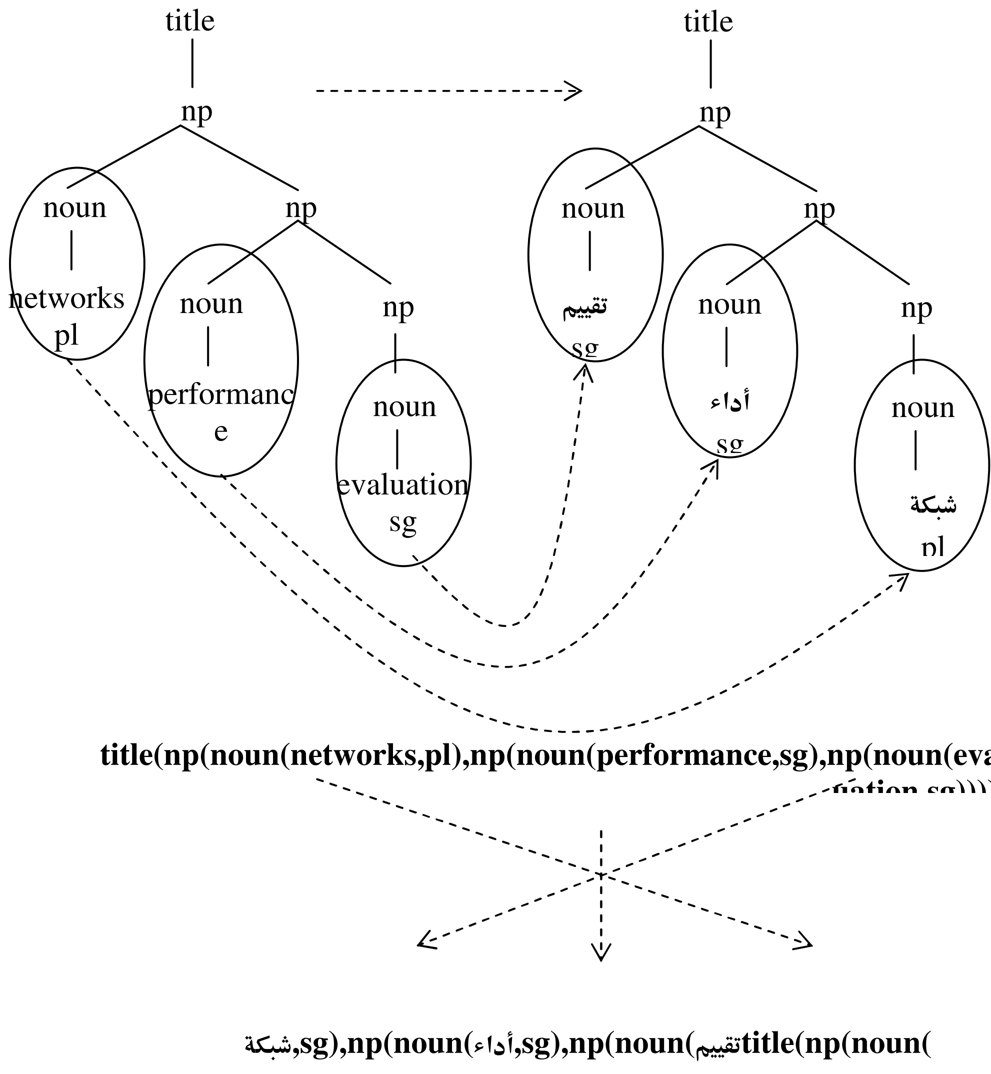


Fig. 3-9 Simple Transfer

In general, almost noun phrases are in a compound form. The translation rule of a compound noun phrase looks as follows.

(2)	[NP:$1, Sept:$2, NP:$3]
	⇔
	[NP:$1, Sept:$2, NP:$3]

As an example, consider the translation of "intelligent search system for bibliographic services" into "نظام بحث ذكي لخدمة ببيلوجرافي". The rule associates $1 with the subtree of "intelligent search system", $2 with the node for "for", and $3 with the subtree for "bibliographic services". Translating each of these then becomes a separate task for transfer, which operates on these subtrees in the same way as on the original tree—attempting to find rules which deal with these sorts of structure. This process is illustrated in Fig. 3-10.

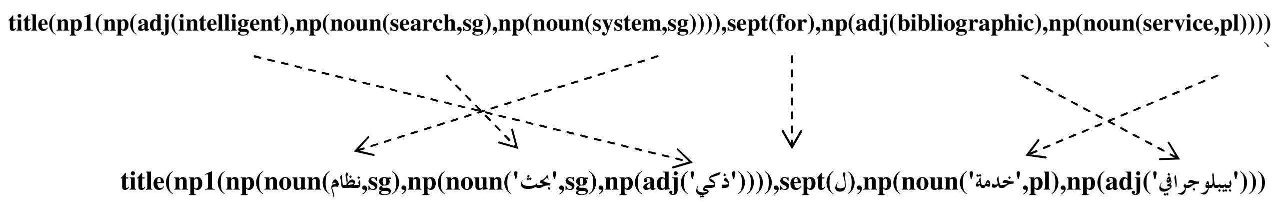


Fig. 3-10 Compound Transfer

In the translation process, lexical mapping is inevitable. In any transfer process, the one-to-zero mapping is undesirable. One-to-zero mapping means the lack of equivalent lexical word in the target language. This phenomena is called "lexical hole" in (Hai, Kawtrakul and Poovorawan, 1997),(Trujillo, 1999). In this work we found this gab when we translate a noun phrase that contains an "of" separator. This translation is described in two steps: restructuring of the English parse tree and reordering of words. The translation rule of a noun phrase that contains the separator "of" looks as follows.

(3) [NP:$1, Sept:of, NP:$2] ⇨[NP:$3[NP:$2, NP:$1]]
⇔
[NP:$3]

As an example, consider the translation of "performance evaluation of routing algorithms" into "تقييم أداء خوارزمية مسارات". The rule associates $1 with the subtree of "performance evaluation" and $2 with the subtree for "routing algorithms". In the first step, the original tree is restructured by dropping the "of" node and switching the arguments of the "of". This yield a tree representation of "routing algorithms performance evaluation" that is to be associated with \$3. In the second step, the normal transfer rules are applied to the tree representation associated with \$3 to get the translation of the original tree. This process is illustrated in Fig. 3.11.

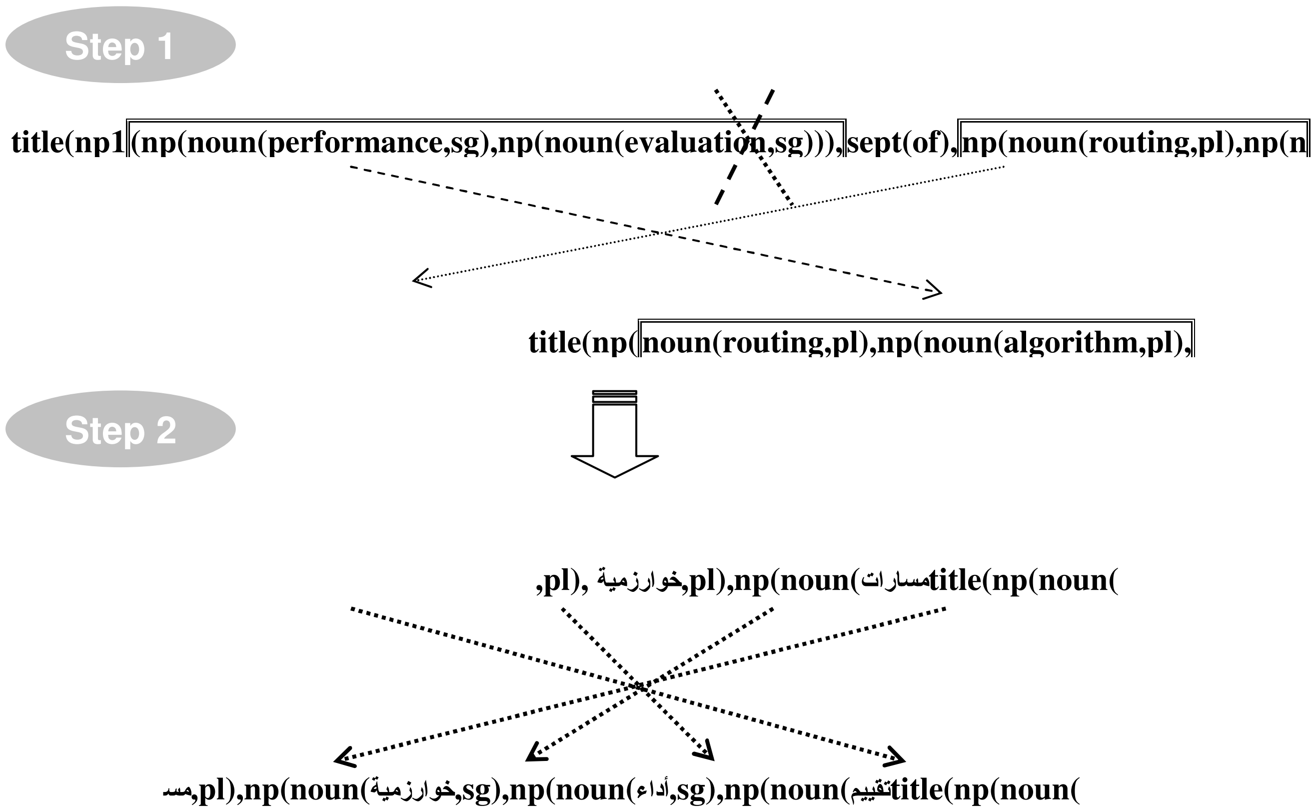


Fig. 3-11 "of" Transfer

The translation of a noun phrase starting in a quantifier is described by a special rule that looks as follows.

(4)	$[w_i:\$1[\text{quantifier form}], w_{i+1}:\$2, \ldots, w_k:\$k]$	$(1 \le i \le k)$
	$\Leftrightarrow$	
	$[w_i:\$1, w_k:\$k, w_{k-1}:\$k-1, \ldots, w_i:\$i]$	$(1 \le i \le k)$

This rule says that the quantifier that fills the head of the noun phrase is translated to an equivalent word that fills the head, i.e. at the same position, of target noun phrase. The rest of the noun phrase is translated in the normal transfer rules. As an example, consider the translation of "comparative study for some software engineering tools" into " دراسة مقارنة لبعض أداة هندسه برمجيات". This process is illustrated in Fig. 3.12.

title(np1(np(noun(comparative,sg),np(noun(study,sg))),sept(for),

title(np1(np(noun(دراسة,sg),np(noun(مقارنة,sg))),sept(ل),
np(quantifier(بعض),np(noun(أداة,pl),np(noun(هندسه,sg),np(noun(برمجيات,pl)))))))

Fig. 3-12 Quantifier Transfer

The translation of a noun phrase starting in a gerund (-ing) form is described by a special rule that looks as follows.

(5)	$[w_i{:}\$1[\text{-ing form}],\ w_{i+1}{:}\$2,\ \ldots,\ w_k{:}\$k]$	*$(1 \le i \le k)$ & $\$1 \notin termbase$*
	$\Leftrightarrow$	
	$[w_i{:}\$1,\ w_k{:}\$k,\ w_{k-1}{:}\$k\text{-}1,\ \ldots,\ w_i{:}\$i]$	*$(1 \le i \le k)$*

This rule says that the gerund form that fills the head of the noun phrase is translated to an equivalent word that fills the head, i.e. at the same position, of target noun phrase. The rest of the noun phrase is translated in the normal transfer rules. As an example, consider the translation of "automating software testing" into "أتمتة اختبار برمجيات". This process is illustrated in Fig. 3.13.

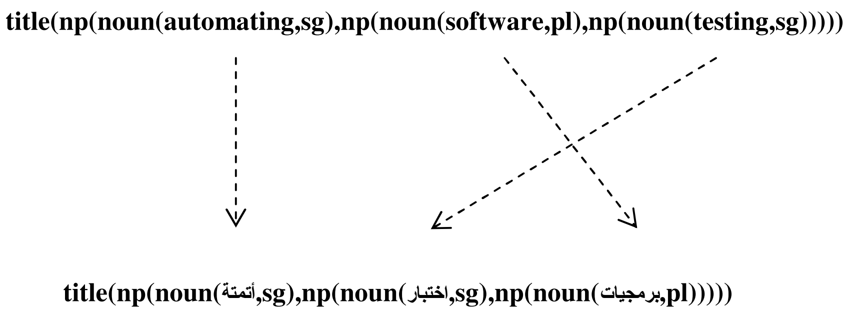


Fig. 3-13 Gerund Transfer

One further condition is added in order to apply this rule. The gerund word that fills the head of the noun phrase must not be a computer term. As an example, consider the translation of "switching techniques for multimedia information transmission over a widearea network" into " تقنيات تحويل لنقل معلومات الوسائط المتعددة علي شبكة متسعة". Since "switching" is a gerund term, the above rule cannot be applied. Special words like the one given have to be kept in a termbase in order to prevent this rule from being applied.

The translation of a noun phrase starting in a word ending in (able) form is described by a special rule that looks as follows.

(6)	[w_i:\$1[-able form], w_{i+1}:\$2, …, w_k:\$k]	*($1 \le i \le k$)*
	⇔	
	[w_i:\$1, w_k:\$k, w_{k-1}:\$k-1, …, w_i:\$i]	*($1 \le i \le k$)*

As an example, consider the translation of "reusable problem solving components" into "اعادة أستخدام مكون حل مشكلة". This process is illustrated in Fig. 3.14.

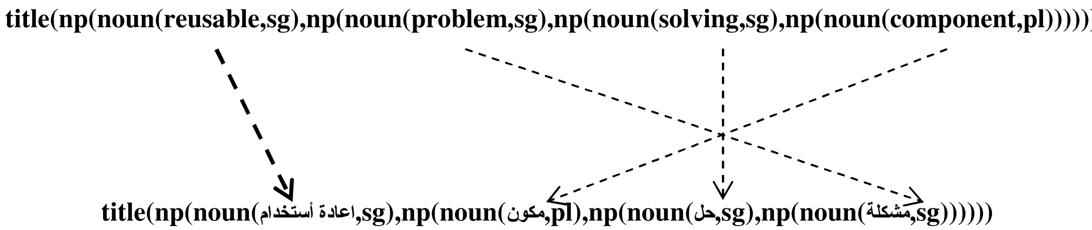


Fig. 3-14 "-able" Transfer

The implementation of the transfer component is shown in Fig.3-15. In this implementation, the special transfer rules are applied first.

```
 transfer(title(NP),title(NP1)):-!,
          transfer(NP,NP1).
transfer(np1(NP,sept(of),NP1),NP6):-
          NP1=np1(NP2,Sept,NP3),!,
          append_np(NP2,NP,NP4),
          NP5=np1(NP4,Sept,NP3),
          transfer(NP5,NP6).

transfer(np1(NP,sept(of),NP1),NP3):-!,
          append_np(NP1,NP,NP2),
          transfer(NP2,NP3).

transfer(np1(NP,sept(ESept),NP1),np1(NP2,sept(ASept),NP3)):-
          dic(ESept,separator,ASept),
          transfer(NP,NP2),
          transfer(NP1,NP3).

transfer(np(quantifier(EW),NP),np(quantifier(AW),NP1)):-
          dic(EW,quantifier,AW),!,
          transfer_np(NP,NP1).

transfer(np(noun(EW,No),NP),np(noun(AW,No),NP1)):-
          same_pos(EW),!,
          dic(EW,noun,AW),
          transfer_np(NP,NP1).
transfer(NP,NP1):-
          transfer_np(NP,NP1).

transfer_np(np(noun(EW,No),NP),NP2):-!,
          dic(EW,noun,AW),
          transfer_np(NP,NP1),
          transfer_np(NP1,noun(AW,No),NP2).
```

```
transfer_np(np(adj(EW),NP),NP2):-
        dic(EW,adj,AW),!,
        transfer_np(NP,NP1),
        transfer_np(NP1,adj(AW),NP2).

transfer_np(np(det(X),NP),np(det(X),NP1)):-!,
        transfer_np(NP,NP1).

transfer_np(np(noun(EW,No)),np(noun(AW,No))):-!,
        dic(EW,noun,AW).

transfer_np(np(Any),Term,E):-!,
        E=np(Any,np(Term)).

transfer_np(np(Any,NP),Term,np(Any,NP1)):-!,
        transfer_np(NP,Term,NP1).

append_np(np(X),NP,np(X,NP)).
append_np(np(X,NP),NP1,np(X,NP2)):-
        append_np(NP,NP1,NP2).
```

Fig. 3-15 Implementation of the Transfer Component

3.3.2 Bi-Lingual Dictionary and the Termbase

In addition to the dictionary a terminological database (also called a "termbase") is provided which is simply just the gerund (–ing) form of a word that corresponds to a terminology. Each entry is a specialized vocabulary used to refer to a terminology of a particular domain that is included in the dictionary. They are identified in order to supplement the existing dictionary for getting a suitable translation of the source text. In our transfer process, these words are needed in order to satisfy condition given in rule (5). Examples of the terminology from computer science are shown in Table 3-5.

Table 3-5 Examples of Termbase entries

modeling	**routing**	**mining**
programming	**switching**	**scheduling**
processing	**reasoning**	**training**
learning	**authoring**	**engineering**

The termbase entry is implemented as a Prolog fact that has the form:

gerund_term(*Term*)

Where *Term* is a term from the source text domain.

Recall that the lexical transfer is done through bilingual dictionary rules. This dictionary relates the source lexical units ("words") to target lexical units. The bi-lingual dictionary entry is represented as prolog facts as follows:

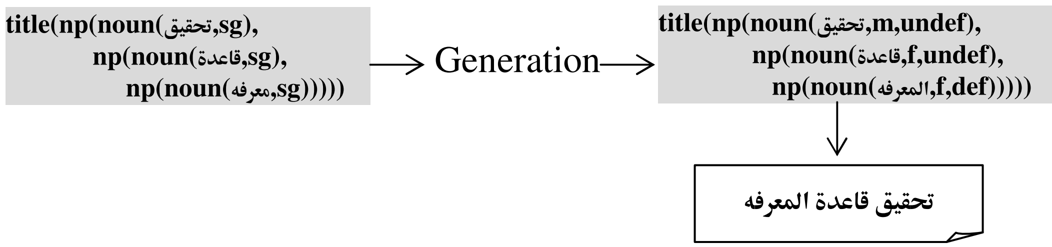
dic(*English_word, Word_Cat, Arabic_word*).

Where:

- The *English_word* is either:
  - A noun in singluar form or an irregular plural
  - An adjective form
  - A separator form
  - A quantifier form
- The *Word_Cat* (word category) is the word noun, adj, separator, or quantifier.
- The *Arabic_word* is the Arabic equivalent of the *English_word*

3.4 The Generation Components

In the third step, sentences are generated by using the target language's parsing rules in reverse, building a sequence of words in the target language that express the meaning in the translated parse tree. The Arabic generator component will synthesis the inflected Arabic word-form based on the morphological features and traverse the syntactic tree to produce the surface Arabic noun phrase. The generation step comprises Arabic dictionary, Arabic morphological synthesizer, and Arabic noun phrase synthesizer.



3.4.1 The Arabic Dictionary

The Arabic dictionary used in this step is a monolingual dictionary containing Arabic morphemes only. Considering nouns and adjectives, they are stored in the dictionary in their singular masculine form. Their synthesis depends on:

- Gender of noun
- Number of noun
- Definition of noun

As is the case with English monolingual dictionary, the irregular forms are explicitly included in the Arabic dictionary.

The following describes the structure of entries for the proposed monolingual Arabic dictionary. Table 3-6 shows examples of these entries.

- *Noun*: a content word which has four features the stem-form, the number, gender, and definition. Nouns are stored in singular form. Irregular plurals have additional entries. Recall that for the translation purposes, we used the stem-form rather than the root-form for nouns.
- *Adjective*: a content word, which has a word-form feature.

Table 3-6 Examples of the monolingual English dictionary entries

Noun					Adjective
Form	***Number***	***Gender***	***Definition***	***Irr. Plural***	***Form***
منهج	sg	m	undef	مناهج	عصبي
مناهج	pl	f	undef	none	طبيعي
مشكلة	sg	f	undef	none	جديد

The dictionary is implemented as Prolog facts. Figure 3.16 shows the implementation of the examples given in Table 3-6.

```
noun('منهج',sg, m,undef, ['مناهج']).     adj('عصبي').
noun('مناهج',pl, f,undef, []).           adj('طبيعي').
noun('مشكلة',sg, f,undef, []).           adj('جديد').
```

Fig.3-16 Implementation of the Arabic dictionary

3.4.2 The Arabic Morphological Synthesizer

The Arabic morphological synthesizer will generate the inflected Arabic word according to the Arabic agreement relationship. This relationship is between words in certain context such that a word in one position follows the word in a corresponding position in some aspects: such as number (single, plural), sex (male, female), and definition (definite, indefinite). The agreement rules are different in different languages. For example, in English, adjective does not agree with its noun in count and sex; we say “good man,” “good girl,” and “good men.”

. The architecture of the morphological synthesizer is given in Figure 3.17. The implementation of the morphological synthesizer is given in Fig. 3.18. This figure shows that the role of the morphological synthesis is to:

- Synthesize defined noun: lookup the noun in the dictionary if it is definite then return; otherwise defines the noun with “ال”.
- Synthesize definite adjective: define the adjective with “ال”.
- Synthesize feminine adjective: feminize the adjective with “ة”.

- Synthesize plural noun:

1. Lookup the stem and features in the dictionary.
2. If the number feature of the dictionary agrees with the number feature of the source noun then return the stem and its features.
3. If the number feature of the source noun is plural and there is an explicit (irregular) plural, return the plural stem and its features.
4. Otherwise, synthesize the regular plural of the Arabic noun (sound masculine plural and sound feminine plural) that agrees with the gender of the stem.

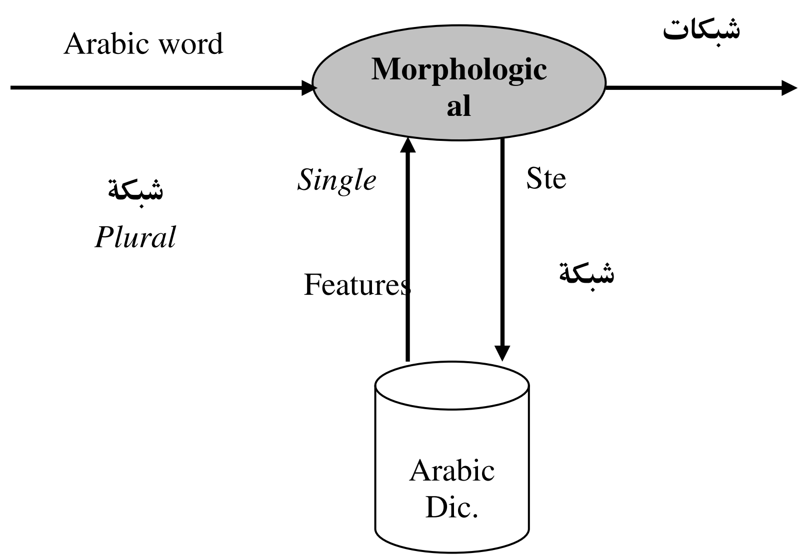


Fig. 3-17 Architecture of the morphological synthesizer

```
synthesize_def_noun(def,W,W):-!.
synthesize_def_noun(_,W,DefW):-
        name(W,W1),
        (append(W2,[32|W3],W1) ->
                append("ال",W2,DefW2),
                append("ال",W3,DefW3),
                append(DefW2,[32|DefW3],DefW1)
        ; append("ال",W1,DefW1)
        ), name(DefW,DefW1).
synthesize_def_adj(W,DefW):-
        name(W,W1),
        ((append(W2,[32|W3],W1),
          name(W4,W2),
          adj(W4)) ->
                append("ال",W2,DefW2),
                append(DefW2,[32|W3],DefW1)
        ; append("ال",W1,DefW1)
        ), name(DefW,DefW1).
synthesize_fem_adj(W,FemW):-
        name(W,W1),
        ((append(W2,[32|W3],W1),
          name(W4,W2),
          adj(W4)) ->
                append(W2,"ة",FemW2),
                append(FemW2,[32|W3],FemW1)
        ; append(W1,"ة",FemW1)
        ), name(FemW,FemW1).
```

Fig.3-18 Implementation of the morphological synthesizer of Nouns and adjectives

3.4.3 The Arabic Noun Phrase synthesizer

The main role of the noun phrase synthesizer is to construct the Arabic noun phrase that represents the translated noun phrase. The architecture of the noun phrase synthesizer is shown in Fig. 3-19. In this architecture, depending on the syntactic category, the word is synthesized by calling the morphological synthesizer.

The noun phrase synthesis consists of three phases. First, the right nouns are built according to the number and definition features. Second, we ensure the agreement relationship between descriptive nouns and adjectives with regard to the sex and definition features. Third, the transformed tree is traversed to produce the final output.

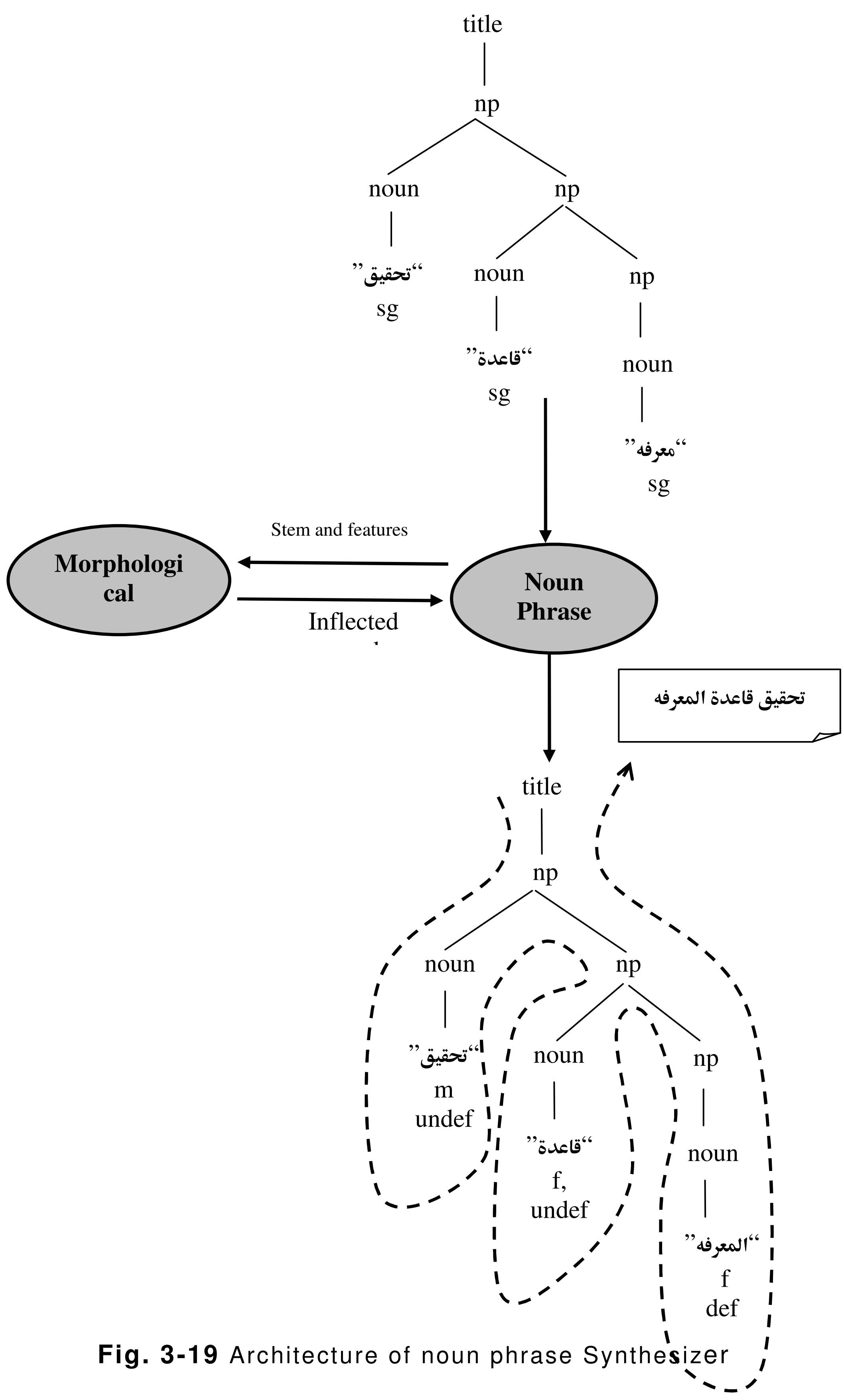


Fig. 3-19 Architecture of noun phrase Synthesizer

This process is shown in Fig. 3.20. In the following we will describe the transformation rules for synthesizing nouns, the transformation for synthesizing adjectives, and the traversal of the tree to produce the final output.

```
synthesis(Tree1):-
        process_noun(Tree1,Tree2),
        process_adj(Tree2,Tree),
        traverse(Tree,L),write_list(L),nl.
```

Fig. 3-20 **Implementation of the Noun Phrase synthesizer**

Transformations of Nouns

A set of transformational grammar rules is applied to synthesize the nouns in the target Arabic noun phrase to get a grammatically correct structure. These transformational rules (Neaama, 1973) are divided into two categories: rules that generate plurals and rules that define nouns with "ال". Since these rules may be applied together, i.e. we may synthesize a defined plural, the implementation of the structure inspection of the noun phrase is put together, see Fig. 3-21. In this implementation, the leave nodes of the tree are transformed according to the lexical features in the Arabic monolingual dictionary and the structural features from applying syntactic rules of Arabic.

```
process_noun(title(NP),title(NP1)):-!,
        process_noun(NP,NP1).
process_noun(np1(NP,Sept,NP1),np1(NP2,Sept,NP3)):-
        process_noun(NP,NP2),
        synthesize_def_genitive_np(Sept,NP1,NP3).
process_noun(np(det(the),NP),NP1):-!,
        synthesize_def_np(NP,NP1).
process_noun(np(det(A),NP),NP1):-
        (A=a;A=an),!,
        synthesize_undef_np(NP,NP1).
process_noun(np(noun(AW,No)),np(noun(AW1,G,D))):-!,
        synthesize_pl_noun(AW,No,AW1,G,D).
process_noun(np(Any),np(Any)):-!.
process_noun(np(noun(M,No),np(noun(MA,No1))),E):-!,
        E=(np(noun(M1,G1,D1),np(noun(MA2,G2,def)))),
        synthesize_pl_noun(M,No,M1,G1,D1),
        synthesize_pl_noun(MA,No1,MA1,G2,D2),
        synthesize_def_noun(D2,MA1,MA2).
process_noun(np(noun(M,No),NP),E):-!,
        E=np(noun(M1,G,D),NP1),
        synthesize_pl_noun(M,No,M1,G,D),
        synthesize_def_np(NP,NP1).
process_noun(np(Any,X1),np(Any,Y1)):-!,
        process_noun(X1,Y1).
```

Fig. 3-21 Implementation of the First Phase of the Nouns Phrase synthesizer

Synthesis of definition

1. Definition with the postfixed noun ("المضاف الية")

The following is the morphological rule that handles the definition of the postfixed noun ("المضاف الية") of a noun phrase in annexation ("تركيب إضافي").

IF the noun phrase is in annexation ("تركيب إضافي")
THEN define the postfixed noun ("المضاف الية")

This rule is the most common rule. The implementation of the definition of the postfixed noun is given in Fig. 3-21. Fig. 3-22 shows an example where the postfixed noun "**معرفه**" is defined

title(np(noun('تحقيق',m,undef),np(noun('قاعدة',f,undef),np(noun('المعرفه',f,def)))))

⇩

title(np(noun('تحقيق',sg),np(noun('قاعدة',sg),np(noun(معرفه,sg)))

Fig. 3-22 Definition with the postfixed noun (“المضاف الية”)

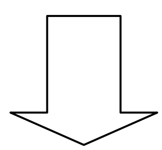
2. *Definition with the article “حرف التعريف ال”*

In English, the article “the” is used to define a noun phrase and the article “a(n)” is used to undefine a noun phrase. In Arabic, the prefix “ال” is used to define a noun in the noun phrase, called definite “معرفه”, whereas its nonexistence would be considered as an indefinite noun, called indefinite “نكرة”. The following is the rule that handles the definition with article.

IF the noun phrase is defined with the article (“حرف التعريف”)
THEN define the noun phrase

An exception to this rule is that proper nouns are not defined with articles, e.g. “مصر” Egypt. Fig. 3-23 shows an example of a definition of a noun with article. In our implementation we differentiate between two cases, see Fig. 3-24. The case where the noun phrase are to be defined and the case where it is to be made undefined. For this purpose, the articles are carried over to the generation phase.

title(np(quantifier(من),np(det(the),np(noun(منهج,pl),np(adj('حديث'))))))



title(np(quantifier(من),np(noun('المناهج',f,def),np(adj('حديث')))))

Fig. 3-23 Definition with the article “حرف التعريف ال”

```
synthesize_def_np(np(noun(MA,No),np(adj(M))),E):-!,
        E=np(noun(MA2,G,def),np(adj(M))),
        synthesize_pl_noun(MA,No,MA1,G,D),
        synthesize_def_noun(D,MA1,MA2).
synthesize_def_np(np(noun(MA,No),NP),E):-
        NP=np(adj(Adj),NP1),!,
        E=np(noun(MA2,G,def),np(adj(Adj),NP2)),
        synthesize_pl_noun(MA,No,MA1,G,D),
        synthesize_def_noun(D,MA1,MA2),
        synthesize_def_np(NP1,NP2).
synthesize_def_np(np(noun(MA,No)),E):-!,
        E=np(noun(MA2,G,def)),
        synthesize_pl_noun(MA,No,MA1,G,D),
        synthesize_def_noun(D,MA1,MA2).
synthesize_def_np(X,Y):-
        process_noun(X,Y).
synthesize_undef_np(np(noun(AW,No)),np(noun(AW1,G,D))):-!,
        synthesize_pl_noun(AW,No,AW1,G,D).
synthesize_undef_np(np(Noun),np(Noun)):-!.
synthesize_undef_np(np(noun(AW,No),Sept,NP),np(noun(AW1,G,D),Sept,NP1)):-!,
        synthesize_pl_noun(AW,No,AW1,G,D),
        process_noun(NP,NP1).
synthesize_undef_np(np(N,Sept,NP),np(N,Sept,NP1)):-!,
        synthesize_def_genitive_np(Sept,NP,NP1).
synthesize_undef_np(np(det(A),NP),NP1):-
        (A=a;A=an),!,
        synthesize_undef_np(NP,NP1).
synthesize_undef_np(np(noun(M,No),NP),np(noun(M1,G,D),NP1)):-!,
        synthesize_pl_noun(M,No,M1,G,D),
        synthesize_undef_np(NP,NP1).
synthesize_undef_np(np(Any,NP),np(Any,NP1)):-!,
        synthesize_undef_np(NP,NP1).
```

Fig. 3-24 Implementation of the definition with article

3. ***Definition*** of a noun in the genitive ("المجرور").

In Arabic, the noun phrase that is preceded by a preposition ("حروف الجر"), called quasi-proposition ("شبه جملة"), would define the noun in genitive ("المجرور"). The following is the rule that handles definition of a noun in genitive.

IF the noun phrase is in quasi-proposition ("شبه جملة")
THEN define the noun in genitive ("المجرور")

In our grammar, the syntactic category of prepositions is separators. Fig. 3-25 shows an example where the preposition "ل" is used to define the plural noun in genitive "خدمات", which result in the defined noun "الخدمات". The implementation of the definition the noun in genitive is given in Fig. 3-26.

title(np1(np(noun(نظام,sg),np(noun(بحث,sg),np(adj(ذكي))))),sept(ل),np(noun(خدمة,pl),np(adj(بيبلوجرافي))))

⇩

title(np1(np(noun(نظام,m,undef),np(noun(البحث,m,def),np(adj(ذكي))))),sept(ل),np(noun(الخدمات,f,def) ,np(adj(بيبلوجرافي)))))))

Fig. 3-25 Definition of a noun in the genitive ("المجرور").

```
synthesize_def_genitive_np(sept(Sept),np(noun(MA,No),Sept1,NP),E):-
        is_genitive(Sept),!,
        E = np(noun(MA2,G,def),Sept1,NP1),
        synthesize_pl_noun(MA,No,MA1,G,D),
        synthesize_def_noun(D,MA1,MA2),
        synthesize_def_genitive_np(Sept1,NP,NP1).
synthesize_def_genitive_np(sept(Sept),NP,NP1):-
        is_genitive(Sept),!,
        synthesize_def_np(NP,NP1).
synthesize_def_genitive_np(_,NP,NP1):-
        process_noun(NP,NP1).

% حروف الجر
is_genitive('من'):-!.
...
```

Fig. 3-26 Implementation of the noun in the genitive

Synthesis of plurals

Recall that the analysis module recognizes the stem of the word and its features. The transfer module maps the stem and features to their Arabic equivalents. In the generation module, from the stem and its features, its plural is synthesized. The following is the rule that handles the generation of plurals.

IF the number feature of the stem is plural
Then generate the plural of the noun

For example, the word "algorithms" is analyzed to the word "algorithm", with the number feature "pl". The transfer maps "algorithm" into

"**خوارزمية**" and its number feature into "pl". In the generation, the sound feminine plural ("جمع المؤنث السالم") of the word, which is "**خوارزميات**", is synthesized because the gender feature of the word is feminine "f". This is shown in Fig. 3-27. The implementation of plurals noun and adjective see Fig. 3-28

```
title(np(noun(تقييم,sg),np(noun(أداء,sg),np(noun(خوارزمية,pl),np(noun(مسارات,pl))))))

title(np(noun(تقييم,m,undef),np(noun(أداء,m,undef),np(noun(خوارزميات,f,undef),np(noun(المسارات,f,def))))))
```

Fig. 3-27 Generation of plurals

```
synthesize_pl_noun(AWord,Numero,AWords,Gend,Def):-
        noun(AWord,Numero,S,D,_),!, AWords=AWord,Gend=S,Def=D.
synthesize_pl_noun(AWord,pl,AWords,Gend,Def):-
        noun(AWord,sg,_,_,[AWords]),!,
        noun(AWords,pl,Gend,Def,[]).
synthesize_pl_noun(AWord,pl,AWords,Gend,Def):-
        noun(AWord,sg,f,D,_),!,
        name(AWord,W1),
        append(W2,"ة",W1),
        append(W2,"ات",W3),
        name(AWords,W3),
        Gend=f,Def=D.
synthesize_pl_noun(AWord,pl,AWords,Gend,Def):-
        noun(AWord,sg,m,D,_),!,
        name(AWord,W1),
        append(W1,"ين",W2),
        name(AWords,W2),
        Gend=m,Def=D.
```

Fig. 3-28 Implementation of generation plurals

Transformations of Adjectives

Agreement between nouns and adjectives

A second set of transformational grammar rules is applied to synthesize the adjectives that agree with their described ("الموصوف") nouns in the target

Arabic noun phrase to get the final grammatically correct structure. These transformational rules are described by the following rules.

IF the noun phrase is in adjective "تركيب وصفي" AND
the described noun ("الموصوف") is defined
THEN define the adjective

IF the noun phrase is in adjective "تركيب وصفي" AND
the described noun ("الموصوف") is feminine
THEN feminize the adjective

Fig. 3-29 shows an example where the adjective "بيبلوجرافي" is defined and feminized to agree with the describe noun "الخدمات". Fig. 3-30 shows the implementation of the second phase of the noun phrase synthesizer. This implementation shows how the adjective inherits the gender and the definition from their described nouns.

title(np(noun(نظام,m,undef),np(noun(البحث,m,def),np(adj(ذكي),sept(ل),np(noun(الخدمات,f,def),
np(adj(بيبلوجرافي)))))))

np(noun(نظام,m,undef),np(noun(البحث,m,def),np(adj(الذكي),sept(ل),np(noun(الخدمات,f,def),
np(adj(البيبلوجرافية)))))))

Fig. 3-29 **Agreement of Nouns and Adjectives**

```
process_adj(title(NP),title(NP1)):-!,
        process_adj(NP,NP1).
process_adj(np(quantifier(Q),NP),np(quantifier(Q),NP1)):-!,
        process_adj(NP,NP1).
process_adj(np(det(D),NP),np(det(D),NP1)):-!,
        process_adj(NP,NP1).
process_adj(np(noun(M,G,D),NP), np(noun(M,G,D),NP1)):-!,
        process_adj(NP,G,D,NP1).
process_adj(np(noun(N,G,D),sept(Sept),NP), np(noun(N,G,D),sept(Sept1),NP1)):-
        adj(Sept),!,
        process_adj(adj(Sept),G,D,adj(Sept1)),
        process_adj(NP,NP1).
process_adj(np(Any,sept(S),NP),np(Any,sept(S),NP1)):-!,
        process_adj(NP,NP1).
process_adj(np(Any),np(Any)):-!.
process_adj(np(adj(Adj),NP),G,DEF, np(adj(Adj1),NP1)):-!,
        process_adj(adj(Adj),G,DEF,adj(Adj1)),
        process_adj(NP,G,DEF,NP1).
process_adj(np(adj(Adj)),G,DEF, np(adj(Adj1))):-!,
        process_adj(adj(Adj),G,DEF,adj(Adj1)).
process_adj(np(adj(Adj),sept(Sept),NP),G,DEF, np(adj(Adj1),sept(Sept1),NP1)):-
        adj(Sept),!,
        process_adj(adj(Adj),G,DEF,adj(Adj1)),
        process_adj(adj(Sept),G,DEF,adj(Sept1)),
        process_adj(NP,NP1).
process_adj(np(adj(Adj),Sept,NP),G,DEF, np(adj(Adj1),Sept,NP1)):-!,
        process_adj(adj(Adj),G,DEF,adj(Adj1)),
        process_adj(NP,NP1).
process_adj(adj(X),G,DEF,E):-!,
        (DEF==def ->
                (G==f ->
                        synthesize_fem_adj(X,X1),
                        synthesize_def_adj(X1,X2)
                ; synthesize_def_adj(X,X2)
                )
        ;       (G==f ->
                        synthesize_fem_adj(X,X2)
                ; X=X2)
        ),
        E=adj(X2).
process_adj(X,_,_,Y):-!,
        process_adj(X,Y).
```

Fig. 3-30 **Implementation of the Second Phase of the Nouns Phrase synthesizer**

Target Noun Phrase Synthesis

The last phase is responsible for polishing and producing the Arabic noun phrase in its right form, See Fig. 3-31. The final parse tree is traversed in a depth-first manner producing an Arabic noun phrase in a list. This list is written in order to get the final output of the noun phrase translation system.

```
traverse(Tree,L):-
        traverse(Tree,L,[]).
traverse(title(NP),L,L0):-!,
        traverse(NP,L,L0).
traverse(np(W,NP),L,L0):-!,
        get_word(W,W1),
        L=[W1|L1],
        traverse(NP,L1,L0).
traverse(np(W,sept('ل'),NP),L,L0):-
        get_word(W,W1),
        traverse(NP,X2,L1,L0),
        get_word(X2,X3),
        name(X3,X4),
        append("ال",X5,X4),!,
        append("لل",X5,X6),
        name(X7,X6),
        L=[W1,X7|L1].
traverse(np(W,sept(S),NP),L,L0):-
        prefix(S),!,
        get_word(W,W1),
        traverse(NP,X2,L1,L0),
        get_word(X2,X3),
        append_atoms(S,X3,X4),
        L=[W1,X4|L1].
traverse(np(W,sept(S),NP),L,L0):-!,
        get_word(W,W1),
        L=[W1,S|L1],
        traverse(NP,L1,L0).
traverse(np(W),L,L0):-!,
        get_word(W,W1),
        L=[W1|L0].
traverse(_,L,L).
traverse(np(X),X,L,L).
traverse(np(X,NP),X,L,L0):- traverse(NP,L,L0).
traverse(np(X,sept(S),NP),X,L,L0):- traverse(S,NP,L,L0).
write_list([]).
write_list([''|T]):-!,
        write_list(T).
write_list([H|T]):-
        write(H), write(' '),
        write_list(T).
get_word(noun(X,_,_),X).
get_word(det(_),'').
get_word(adj(X),X).
get_word(quantifier(X),X).
append_atoms(A,B,C):-!,
        name(A,A1),name(B,B1),append(A1,B1,C1),
        name(C,C1).
```

Fig. 3-31 Implementation of the Third Phase of the Nouns Phrase synthesizer

CHAPTER **4**

EVALUATION OF MACHINE TRANSLATION SYSTEMS

The evaluation of MT systems is an important issue, which has received considerable attention. There are many reasons for evaluating MT systems (Trujillo 1999). For example, one might want to decide whether the quality of the output is sufficiently high for a particular purpose. There is no standard approach to evaluation, and each application of MT may require its own distinct evaluation methodology. In this chapter we developed an evaluation methodology to evaluate the translation system and applied this methodology on three experiments. The experiments result are analyzed and discussion.

4.1 Overview

First attempts in the field of MT evaluation (as reported in [Woo 94]) were based on some false hypotheses. One of them can described as follow: If you translate a text from Italian into French, for instance, and then back to Italian, you will get an exact copy of the original text – provided that your translation system worked well. And every difference between the original and the translated Italian text will indicate an error your system has made. This hypothesis is not correct –consider that even an extraordinarily good human translation will not reach this goal. Beside, to translate in such a manner is not the aim of the translation anyway.

A second hypothesis is based on the same misconception. You choose a text to have it translated by a machine and also by a good human translator, and then you compare the output. You are tempted to say: the closer the results, the better the machine system. This view is wrong again, for similar reasons. It is possible to translate a sentence in two different ways and to get two different and nevertheless correct results. The machine-translated output can be of the same quality although only similar to the human translation.

Of course, this does not mean that the qualified achievements of human translators have lost their status as a measure for any kind of translation. It is still desirable to reach the human translation level. As soon as we are no longer able to decide whether a text has been translated by a machine or by a person, we have reached our goal. We compare this to the situation in the Turing test. Turing proposed to define the intelligence of a computer according to the following setup: You sit on a computer and enter some questions, which are answered immediately. If you cannot decide whether the answers are given by the computer itself or by a person sitting in front of a terminal next door, the system shows intelligent behavior.

One way to evaluate a MT system, which avoid the problems mentioned, is to check whether the target text still contains the information of the original. For this purpose, you may give a target text to persons who do not know the source language and ask them questions about the text. These questions must be chosen in a way that they can only be answered with information from the text. In this way you can gain knowledge of how much of the original information is preserved in the translated text.

Another tricky problem in the evaluation process is the judgement of translated texts by experts. Whatever you do, you will get individual and subjective results. Our texts have shown that is possible to flatten these discrepancies by asking the participants for a relative ranking of some target texts rather than for an absolute rating.

As reported in the literature (e.g. in [WOO94]), you should not compare different language pairs or different translation directions. Otherwise, you will not be able to interpret your result since you do not have a common ground for a valid comparison.

Other reports ([Jdb93],[Fla94]) have shown that it is possible to look at errors that appear in a translated text and to classify them. You can examine to what extent an error affects the understanding of a text or, how easy it is to correct the error. It is more difficult, if sometimes not impossible, to find out why the translation system has made a certain error, Such diagnoses are often purely hypothetical. Nevertheless, error classification seems to be a useful tool for MT evaluation.

4.2 The Evaluation Methodology

In order to evaluate the correctness of our MT system we developed an evaluation methodology. This methodology is based on a comparison between the system output with the original translation of the input text. The following steps describe the evaluation methodology:

1. Run the system on the selected test case.
2. Compare the original translation with the system output.
3. Classify the problems that arise from the mismatches between the two translations.

4. Assign a suitable score for each problem. A range of score between 0 and 10 determines the correctness of the translation. While 0 indicates absolutely incorrect translation and 10 indicates absolutely correct (matched) translation.
5. When a situation belongs to multiple problems compute its score average.
6. Determine the correctness of the test case by computing the percentage of the total scores.
7. Suggest possible solutions for the identified problems and apply the necessary improvements to the MT system.

In order to improve the translation output the evaluation methodology is applied on successive stages that includes a cycle of translation, error identification, correction, and re-translation until no more changes can be made cost effectively and without diminishing the quality of other areas of the system. In the following subsections we describe three successive experiments that evaluates the system and incrementally improves its output.

4.2.1 Experiment I

The purpose of this experiment is to investigate whether the machine translation system is sufficiently robust for translating English noun phrases to Arabic. The evaluation methodology is applied on 66 real thesis titles from computer science domain. This translation system gives 15 matched translations, See table 4-1.

Table 4-1 Matched translation in Experiment I

Titles	ترجمة آلية	ترجمة أصلية
Diagnostic method using non_monotonic reasoning	طريقة تشخيصيه باستخدام سببيه غير رتيبة	طريقة تشخيصيه باستخدام سببيه غير رتيبة
Cryptography for computer networks security	تشفير البيانات لتأمين شبكات الحاسب	تشفير البيانات لتأمين شبكاتالحاسب
An object_oriented framework for accounting systems	إطار شيئي التوجه لنظم المحاسبية	إطار شيئي التوجه لنظم المحاسبية
A tool for knowledge discovery in life insurance data	أداة لاكتشاف المعرفة في بيانات تأمين الحياة	أداة لاكتشاف المعرفة في بيانات تأمين الحياة
Developing a text retrieval system with a natural language interface	تطوير نظام استرجاع نص مع واجهة لغة طبيعية	تطوير نظام استرجاع نص مع واجهة لغة طبيعية
Data mining for improving data bases capabilities	تنقيب البيانات لتحسين قدرات قواعد البيانات	تنقيب البيانات لتحسين قدرات قواعد البيانات
Diagnostic reasoning method using dynamic model	طريقة الاستنتاج التشخيصي باستخدام نموذج دينا مكي	طريقة الاستنتاج التشخيصي باستخدام نموذج دينا مكي
Authorization system in object_oriented data bases	نظام الصلاحية في قواعد البيانات الشيئية التوجه	نظام الصلاحية في قواعد البيانات الشيئية التوجه
Towards an object_oriented data bases	نحو قاعدة بيانات شيئية التوجه	نحو قاعدة بيانات شيئية التوجه
Intelligent dynamic scheduling for real time systems	جدولة دنياميكية ذكية لنظم الوقت الواقعي	جدولة دنياميكية ذكية لنظم الوقت الواقعي
On fuzzy activity network analysis	من تحليل شبكة النشاط المشوش	من تحليل شبكة النشاط المشوش
A firewalls based scheme for computer networks security	طريقة قائمة علي حوائط نارية لتأمين شبكات الحاسب	طريقة قائمة علي حوائط نارية لتأمين شبكات الحاسب
Learning from analogical examples	تعلم من الأمثلة الشبيهة	التعلم من الأمثلة الشبيهة
A tool for local cash flow forecasting based on fuzzy neural network	أداة لتنبؤ التدفق المحلي القائم علي الشبكات العصبية المشوشة	أداة لتنبؤ التدفق المحلي القائم علي الشبكات العصبية المشوشة
Distributed data bases system: features and impact on applications	نظم قواعد البيانات الموزعة : خصائص و تأثير علي التطبيقات	نظم قواعد البيانات الموزعة : الخصائص و التأثير علي التطبيقات

The remaining 51 titles have problems that arise from the divergences and mismatches between source and target phrases. The following classifies these problems and assigns suitable scores for them.

1. ***Synonyms of a noun.*** This problem appeared because different synonyms of a noun are used. We give an output that belongs to this problem 10.

2. ***Definition with the article***. This problem appeared because the noun phrase that are preceded by “a(n)” is translated as if it were preceded by "the". In other words, the translation nouns and adjectives of this noun phrase are defined. We give an output that belongs to this problem 9.
3. ***Gender of the adjective.*** In general, the adjective agrees with the described noun that precedes it. This problem appeared because the adjective is feminized even though the noun that precedes it has the gender male. This means that the described noun is not the one that precedes the adjective. We give an output that belongs to this problem 9.
4. *Order of the adjective*. This problem appeared because the translation of the adjective relative to its described noun is not translated in its right order. In other words, the adjective does not follow the described noun in order. We give an output that belongs to this problem 9.
5. *Order of simple noun phrases.* Recall that a compound noun phrase is two or more noun phrases separated by a separator. This problem appeared because the order of the translation of simple noun phrases is different. We give an output that belongs to this problem 9.
6. *Successive words form an expression*. This problem appeared because the successive words that form an expression is translated separately. We give an output that belongs to this problem 9.
7. ***Translation of a preposition***. This problem appeared because the translation differs in the translation of a preposition. We give an output that belongs to this problem 8.
8. ***Conjunction with "and".*** In general, the conjunction "and" separates noun phrases. This problem appeared because in some situations an exception is made to use the "and" to separate two nouns. We give an output that belongs to this problem 8.

9. *Adding the preposition "ل"*. This problem appeared because the translation differs in adding the preposition "ل" in the translation. This means that we break down a simple noun phrase in two noun phrases separated by this preposition. We give an output that belongs to this problem 7.

10. ***Translation with additional words.*** This problem appeared because the original translation contains extra words that have no corresponding words in the input noun phrase. We give an output that belongs to this problem 7.

Table 4-2 Evaluation results of Experiment I

Title	Machine Translation	Original Translation	Problem No.	Score
A concurrent constraint logic programming language	لغة برمجة منطقية مقيدة متزامنة	لغة للبرمجة المنطقية المقيدة المتزامنة	1	9
Project resource allocation with the support of an expert system based on fuzzy logic	تحديد مورد المشروع مع دعم نظام خبير قائم على المنطق المشوش	تخصيص موارد المشروع بمساعدة نظام خبير مبني علي المنطق المشوش	1	10
Towards a knowledge modeling approach in risk analysis domain	نحو منهج نمذجة معرفة في مجال تحليل المخاطر	نحو أسلوب لنمذجة المعرفة في مجال تحليل المخاطر	3	8.7
A computer based clinical information system	نظام معلومات إكلينيكية قائمة على حاسب	نظام معلومات إكلينيكي باستخدام الحاسب الآلي	4	8.75
Automatic generation of explanation for expert systems	توليد أتوماتيكي التوضيح لنظم خبيرة	توليد أتوماتيكي للتوضيح في النظم خبيرة	2	7.5
Enhancing authoring systems with the knowledge base concepts	تحسين نظم التأليف مع مفاهيم قاعدة المعرفة	تحسين نظم التأليف بمفاهيم قاعدة المعرفة	1	9
Quality assurance for information system development	تأكيد الجودة لتطوير نظم المعلومات	تأكيد الجودة في تطوير نظم المعلومات	1	9
Comprehensive study on neural networks applications in mathematical programming	دراسة شاملة علي تطبيقات الشبكات العصبية في البرمجة الرياضية	دراسة شاملة لتطبيقات الشبكات العصبية في البرمجة الرياضية	1	9
Goal programming within the frame of constraint logic programming	برمجة هدفيه من خلال إطار البرمجة المنطقية المقيدة	البرمجة الهدفيه في إطار البرمجة المنطقية المقيدة	1	9
Behavioral simulation of protein based digital systems	محاكاة سلوك النظم الرقمية القائمة علي البروتين	محاكاة سلوك النظم الرقمية المبنية بالبروتين	1	10
A knowledge based system for software maintenance	نظام قائم علي معرفة لصيانة البرمجيات	نظام مبني علي قواعد المعرفة لصيانة البرامج	3	8.7
Multimedia based educational software	برمجيات تعليمية قائمة علي وسائط متعددة	برمجيات تعليمية باستخدام الوسائط المتعددة	1	10
An authoring shell for intelligent tutoring systems	أداة تأليف لنظم التعليم الذكي	أداة تأليف لنظم التعليم الذكية	1	9
Developing a cost function in a competitive market using an operation_research technique	تطوير دالة تكلفة في سوق تنافسي باستخدام أساليب بحوث العمليات	تطوير دالة تكلفة في سوق تنافسية باستخدام أساليب بحوث العمليات	1	9

Title	**Machine Translation**	**Original Translation**	**Problem No.**	**Score**
Building a decision support system for organization in competitive market	بناء نظام دعم قرار لمنظمات في السوق التنافسي	بناء نظام لدعم اتخاذ القرار للمنظمات في سوق تنافسية	1	7
A simulation model for transient storage stations	نموذج محاكاة لمحطات التخزين المؤقت	نموذج محاكاة لمحطات التخزين المؤقتة	1	9
An object_oriented approach for building hypermedia systems	منهج شيئي التوجه لبناء نظم الوسائط المترابطة	أسلوب شيئي التوجه لبناء نظم الوسائط المترابطة	1	10
Automatic knowledge acquisition tool for scheduling systems	أداة اكتساب المعرفة الأتوماتيكية لنظم الجدولة	أداة لاستخلاص المعرفة آليا انظم الجدولة	2	8.5
Call admission control in high speed computer networks	مراقبة أجراء الاتصال في شبكات الحاسب السر يع العالي	التحكم في أجراء الاتصالات خلال شبكات الحاسبات ذات السر عات العالية	3	8.7
Evaluation and improvement of work efficiency in wastewater stations using simulation technique	تقييم و تطوير كفاءة العمل في محطات الصرف الصحي باستخدام تقنية المحاكاة	تقييم و تطوير كفاءة العمل بمحطات الصرف الصحي باستخدام أسلوب المحاكاة	2	9.5
On fuzzy multiobjective transportation problem	من مشكلة نقل المتعددة الأهداف المشوشة	في مشكلة النقل المشوشة متعددة الأهداف	1	9
An expert system for engineering insurance	نظام خبير لتأمين الهندسة	نظام خبير في مجال التأمين الهندسي	3	8.3
A computer simulation model for analyzing the accounting systems for industrial companies	نموذج محاكاة حاسب لتحليل النظم المحاسبية للمنشآت الصناعية	نموذج محاكاة الكمبيوتر لتحليل نظام المحاسبية للمنشآت الصناعية	2	9.5
A hybrid framework for optimal system design	إطار هجين لتصميم النظام الأمثل	هيكل هجين للتصميم الأمثل للنظم	2	9.5
Tools for developing multimedia educational systems	أدوات لتطوير النظم التعليمية الوسائط المتعددة	أدوات لبناء النظم التعليمية متعددة الوسائط	1	10
Intelligent dynamic scheduling for real time systems	جدولة ديناميكية ذكية لنظم الوقت الواقعي	الجدولة الديناميكية الذكية لأنظمة الوقت الحقيقي	1	10
An interactive fuzzy goal programming model for multiobjective decision support systems	نموذج برمجة هدفيه مشوشة تفاعلية لنظم دعم قرارات المتعددة الأهداف	نموذج برمجة هدفيه مشوشة تفاعلية لنظام اتخاذ القرارات المتعددة الأهداف	1	10
A computer tool for knowledge acquisition in different domains	أداة حاسب لاكتساب المعرفة في المجالات المختلفة	أداة حاسوبية لاستخلاص المعرفة في المجالات المختلفة	1	10

Title	**Machine Translation**	**Original Translation**	**Problem No.**	**Score**
Agent approach for knowledge base systems	منهج العميل لنظم قاعدة المعرفة	أسلوب العميل لنظم قواعد المعرفة	1	10
Neural networks approach for controlling the operation of power transformer differential relay	منهج الشبكات العصبية لتحكم في تشغيل المتم الفرقي محول القدرة	التحكم في تشغيل المتم الفرقي لمحولات القدرة بأسلوب الشبكات العصبية	2	8
Expert system development tool based on routine design generic task	أداة تطوير النظام الخبير القائم علي المهمة العامة التصميم الروتيني	تطوير أداة لبناء النظم الخبيرة مبنية علي البرنامج العام للتصميم الروتيني	3	8
Investigating a new constraint solving approach within constraint logic programming	بحث منهج حل قيد جديد من خلال البرمجة المنطقية المقيدة	بحث طريقة جديدة لحل القيود من خلال البرمجة المنطقية المقيدة	2	7
Performance evaluation of proposed routing techniques for the internet	تقييم أداء تقنيات المسارات المقترحة للإنترنت	تقييم أداء أساليب مقترحة لاختيار المسارات لشبكة الإنترنت	2	8
A study on recent approaches in goal programming with an application	دراسة علي مناهج حديثة في البرمجة الهدفيه مع تطبيق	دراسة للأساليب الحديثة في البرمجة الهدفيه في حالة تطبيقية	4	8.75
Visual composition for intelligent tutoring systems	تجمع مرئي لنظم التعليم الذكي	التجمع المرئي للأنظمة التعليم الذكية	1	9
Integrating artificial intelligence with project management scheme	تكامل الذكاء الاصطناعي مع طريقة إدارة المشروع	التكامل بين الذكاء الاصطناعي و إدارة المشروعات	2	8
An integer goal programming model for improving athletics training schedule	نموذج برمجة هدفية صحيحة لتحسين جدول التدريب الرياضي	نموذج برمجة الاهداف الصحيحة لتحسين جداول برامج تدريب رياضة العاب القوي	1	8
Arabic text retrieval and classification using neural network	استرجاع نص اللغة العربية و تصنيف باستخدام الشبكة العصبية	استرجاع وتصنيف النصوص العربية باستخدام الشبكات العصبية	1	8
A computer_assisted tool for evaluation of Pascal programs	أداة بمساعدة حاسب لتقييم برامج الباسكال	أداة لتقييم برامج بلغة الباسكال بمساعدة الحاسب	1	9
Computer simulation for a hospital resources planning	محاكاة الحاسب لتخطيط موارد مستشفي	أسلوب المحاكاة باستخدام الحاسب الآلي لتخطيط موارد مستشفي	1	7
Decision support system for fuzzy multiple projects resources scheduling	نظام دعم القرار لجدولة موارد المشروعات المتعددة المشوشة	نظام دعم القرارات لجدولة الموارد المشوشة للمشروعات المتعددة	1	9

Title	**Machine Translation**	**Original Translation**	**Problem No.**	**Score**
Knowledge modeling and representation using logic programming and constraint based approach	نمذجة المعرفة وتمثيل باستخدام البرمجة المنطقية و منهج قائم علي قيود	نمذجة وتمثيل المعرفة باستخدام البرمجة المنطقية و أسلوب مبني علي القيود	2	9
Evaluation of flow control in integrated service digital network	تقييم مراقبة التدفق في الشبكة الرقمية الخدمية المتكاملة	تقييم سياسات التحكم في معدلات تدفق البيانات في شبكات الحاسبات الرقمية متكاملة الخدمات	1	7
Computer based system for item banks : analysis , design and implementation	نظام قائم علي حاسب لبنوك الأسئلة : تحليل و تصميم وتنفيذ	تحليل و تصميم و /نظام آلي لبنوك الأسئلة تنفيذ	1	7
Assessment and applicability of software maintenance methodologies	تقييم و إ مكانية تطبيق مناهج صيانة البرمجيات	تقييم منهجيات صيانة البرمجيات و إمكانية تطبيقها	1	8
Developing a tool for the fusion method in object_oriented programming	تطوير أداة لطريقة الدمج في البرمجة الشيئية التوجه	تطوير أداة لاستخدامها في طريقة الدمج في برمجة شيئية التوجه	1	7
Expert system development tool based on hierarchical classification generic task	أداة تطوير النظام الخبير القائم على المهمة العامة التصنيف الهرمي	تطوير أداة لبناء النظم الخبيرة مبنية على طريقة البرنامج العام للتصنيف الهرمي	2	8.5
Decision support system for inventory management with an application	نظام دعم القرار لأدارة المخزون مع تطبيق	نظام دعم اتخاذ القرار لأدارة المخزون مع حالة تطبيقية	2	8.5
Methodologies for expert system development	مناهج لتطوير النظام الخبير	طرق منهجية لتطوير النظم الخبيرة	1	7
Analysis of data mining methodologies	تحليل مناهج تنقيب البيانات	تحليل منهجيات التنقيب عن البيانات	1	7
Design of an oblique decision tree classifier using fuzzy measures	تصميم مصنف شجرة قرار مائل باستخدام القياسات المشوشة	تصميم مصنف لشجرة قرارات مائلة باستخدام القياسات المشوشة	2	8
Total				431

The evaluation results are shown in Table 4-2. This table shows that the 51 noun phrases take 431 total score. The correctness of the 66 titles is about 88% in average

Improving the Translation Output

The analysis of the problems given above establishes that improving the translation is possible. One possibility of improvement is by acquiring more rules. These rules will be extracted from analyzing the three-phase translation process described in Chapter 3. Another possibility is by relying on the backtracking regime of Prolog that can be used for producing non-deterministic output. This feature will give all possible translations for the input text and let the user choose the suitable correct translation from alternatives. In following paragraphs we will explain the proposed solution, if any, for each problem.

Synonyms of a noun

This problem will be left as is since the different use of synonyms does not affect on the translation. This is explained by the following example.

Title	Machine translation	Original translation
An object_oriented approach for building hypermedia systems	منهج شيئي التوجه لبناء نظم وسائط مترابطة	أسلوب شيئي التوجه لبناء نظم الوسائط المترابطة

Definition with the article

This problem leads to a paradox and cannot be dealt with because we were not able to formulate a comparative grammar rule nor a generation rule that can map the undefined English noun phrase to a defined Arabic noun phrase. This is explained by the following example.

Title	Machine translation	Original translation
A concurrent constraint logic programming language	لغة برمجة منطقية مقيدة متزامنة	لغة للبرمجة المنطقية المقيدة المتزامنة

Gender of the adjective

This problem is solved by making an agreement between the adjective of the first noun in the noun phrase. This is specified by the following generation rule.

IF the noun phrase is in adjective "تركيب وصفي" AND
the noun phrase contain more than one noun
before the adjective
THEN apply the transformations of adjectives as if the first noun
is the describe noun("الموصوف")

The following example shows the improvement realized by the addition of this rule.

Title	Machine translation	Original translation
An authoring shell for intelligent tutoring systems	أداة تأليف لنظم التعليم الذكية	أداة تأليف لنظم التعليم الذكية

Order of the adjective

This problem is solved by switching the adjectives that are followed by two nouns with the first noun, then adding the preposition "ل" to the first noun. This is specified by the following transfer rule.

```
          [ADJ1:$1, N1:$2, N2:$3]
(7)     ⇨ [N1:$2,ADJ1:$1,N2:$3]
        ⇔ [ N2:$3, ADJ1:$1, Sept:'ل', N1:$2 ]
```

This rule includes two steps. The first step is switching the adjectives that are followed by two nouns with the first noun. This means the adjective is describing the second noun. For example "proposed routing techniques" it will be transform to "routing proposed techniques" and in the second step it will be add the preposition "ل" The following example shows the improvement realized by the addition of this rule.

Title	Machine translation	Original translation
Performance evaluation of proposed routing techniques for the internet	تقييم أداء تقنيات مقترحة للمسارات لشبكة الإنترنت	تقييم أداء أساليب مقترحة لاختيار المسارات لشبكة الإنترنت

Order of simple noun phrases

This problem will be left as is since the order of noun phrases does not affect the translation. This is explained by the following example.

Title	Machine translation	Original translation
Neural networks approach for controlling the operation of power transformer differential relay	منهج الشبكات العصبية لتحكم في تشغيل المتم الفرقي محول القدرة	التحكم في تشغيل المتم الفرقي لمحولات القدرة بأسلوب الشبكات العصبية

Successive words form an expression

This problem is solved by adding a compound form such as "constraint logic programming" ("البرمجة المنطقية المقيدة") of the successive words to the dictionary.

Translation of a preposition

Adding the different translations of the preposition to the dictionary solves this problem. The non-deterministic feature of the system would produce alternative translations. For example, the preposition for is translated either as a "ل" or "ق". This is explained by the following example.

Title	Machine translation	Original translation
Quality assurance for information system development	تأكيد الجودة في تطوير نظم المعلومات	تأكيد الجودة في تطوير نظم المعلومات

Conjunction with "and"

This problem is solved by adding the following grammar rule to deal with the exception made when "and" separates two nouns.

np(np(Noun1,Sept,Noun2,NP)) → noun(Noun1), sept(Sept), noun(Noun2),np(NP) ,{Sept=sept(and)}.

This also leads to adding the following transfer rule.

(8) [NP:$1,N1:$2,Sept:and,N2:$3]
⇔ [NP:$4[N1:$2,Sept:and, N2:$3],NP:$1]

The following example shows the improvement realized by the addition of these rules.

Title	Machine translation	Original translation
Knowledge modeling and representation using logic programming and constraint based approach	نمذجة وتمثيل المعرفة باستخدام البرمجة المنطقية و أسلوب مبني علي القيود	نمدحه و تمثيل المعرفة باستخدام البرمجة المنطقية و منهج قائم علي القيود

Adding the preposition "ل"

This problem is solved by adding "ل" for specific patterns. Four situations were identified to fill this gab. Consequently, four transfer rules were added. These are described as follows:

1. Adding the preposition "ل" to the first noun in the pattern "adjective+noun+adjective+noun". This is specified by the following transfer rule.

(9) [ADJ1:$1, N1:$2, ADJ2:$3, N2:$4] ⇔
[N2:$4, ADJ2:$3, Sept:'ل', N1:$2, ADJ1:$1]

The following example shows the improvement realized by the addition of these rules.

Title	Machine translation	Original translation
Expert system development tool based on routine design generic task	أداة تطوير النظام الخبير القائم علي المهمة العامة للتصميم الروتيني	تطوير أداة لبناء النظم الخبيرة مبنية علي البرنامج العام للتصميم الروتيني

2. Adding the preposition "ل" to the second noun that occurs in an adjective that are followed by three nouns. This is specified by the following transfer rule.

(10) [ADJ1:$1, N1:$2, N2:$3, N3:$4] ⇔
[N3:$4, Sept:'ل', N2:$3, N1:$2, ADJ1:$1]

The following example shows the improvement realized by the addition of this rule.

Title	Machine translation	Original translation
Design of an oblique decision tree classifier using fuzzy measures	تصميم مصنف لشجرة قرار مائلة باستخدام القياسات المشوشة	تصميم مصنف لشجرة قرارات مائلة باستخدام القياسات المشوشة

3. Adding the preposition "ل" to the second noun that occurs in two noun followed by an adjective and noun. This is specified by the following transfer rule.

(11) [N1:$1, N2:$2, ADJ:$3, N3:$4] ⇔
[N3:$4, ADJ:$3, Sept:'ل', N2:$2, N1:$1]

The following example shows the improvement realized by the addition of this rule.

Titles	Machine translation	Original translation
Neural networks approach for controlling the operation of power transformer differential relay	منهج الشبكات العصبية لتحكم في تشغيل المتمم الفرقي لمحول القدرة	التحكم في تشغيل المتم الفرقي لمحولات القدرة بأسلوب الشبكات العصبية

4. Adding the preposition "ل" when only a noun occurs after the preposition "of". This is specified by the following transfer rule.

(12) [NP:$1, Sept:of, N:$2] ⇔ [NP:$1,Sept:'ل', N:$2]

The following example shows the improvement realized by the addition of this rule.

Titles	Machine translation	Original translation
Automatic generation of explanation for expert systems	توليد أتوماتيكي للتوضيح في النظم الخبيرة	توليد أتوماتيكي للتوضيح في النظم الخبيرة

Translation with additional words

This problem will be left as is since the addition of the new words cannot be expected. This is explained by the following example.

Title	Machine translation	Original translation
Computer simulation for a hospital resources planning	محاكاة باستخدام الحاسب لتخطيط موارد مستشفي	أسلوب المحاكاة باستخدام الحاسب الآلي لتخطيط موارد مستشفي

4.2.2 Experiment II

The purpose of this experiment is to test the system after implementing the solutions to problems encountered in experiment I. So the system is run again on the same test set. The obtained results along with the number of possible translations are shown in Table 4-3.

Table 4-3 **Evaluation result of Experiment II**

	Titles	Frequency
1.	**A concurrent constraint logic programming language**	1
2.	Project resource allocation with the support of an expert system based on fuzzy logic	2
3.	Towards a knowledge modeling approach in risk analysis domain	4
4.	A computer based clinical information system	2
5.	Automatic generation of explanation for expert systems	2
6.	Enhancing authoring systems with the knowledge base concepts	2
7.	Quality assurance for information system development	2
8.	Comprehensive study on neural networks applications in mathematical programming	2
9.	Goal programming within the frame of constraint logic programming	4
10.	**Behavioral simulation of protein based digital systems**	1
11.	**A knowledge based system for software maintenance**	2
12.	**Multimedia based educational software**	1
13.	Developing a cost function in a competitive market using an operation_research technique	1

	Titles	Frequency
14.	**Building a decision support system for organization in competitive market**	2
15.	**A simulation model for transient storage stations**	4
16.	**An object_oriented approach for building hypermedia systems**	2
17.	**Automatic knowledge acquisition tool for scheduling systems**	4
18.	**Call admission control in high speed computer networks**	2
19.	**Evaluation and improvement of work efficiency in wastewater stations using simulation technique**	1
20.	**On fuzzy multiobjective transportation problem**	1
21.	**An expert system for engineering insurance**	2
22.	**A computer simulation model for analyzing the accounting systems of industrial companies**	2
23.	**A hybrid framework for optimal system design**	2
24.	**Tools for developing multimedia educational systems**	4
25.	**Intelligent dynamic scheduling for real time systems**	4
26.	**An interactive fuzzy goal programming model for multiobjective decision support systems**	8
27.	**A computer tool for knowledge acquisition in different domains**	2
28.	**Agent approach for knowledge base systems**	2
29.	**Neural networks approach for controlling the operation of power transformer differential relay**	12
30.	**Expert system development tool based on routine design generic task**	4
31.	**Investigating a new constraint solving approach within constraint logic programming**	4
32.	**Performance evaluation of proposed routing techniques for the internet**	16
33.	**A study on recent approaches in goal programming with an application**	8
34.	**Visual composition for intelligent tutoring systems**	4
35.	**Distributed data bases system features and impact on applications**	6
36.	**Integrating artificial intelligence with project management scheme**	2
37.	**An integer goal programming model for improving athletics training schedule**	8
38.	**Arabic text retrieval and classification using neural network**	1
39.	**A computer_assisted tool for evaluation of Pascal programs**	2
40.	**Computer simulation for a hospital resources planning**	2
41.	**Decision support system for fuzzy multiple projects resources scheduling**	2
42.	**Knowledge modeling and representation using logic programming and constraint based approach**	1
43.	**Evaluation of flow control in integrated service digital network**	1
44.	**Computer based system for item banks : analysis , design and implementation**	2
45.	**Assessment and applicability of software maintenance methodologies**	1
46.	**Developing a tool for the fusion method in object_oriented programming**	2
47.	**Expert system development tool based on hierarchical classification**	4

	Titles	Frequency
	generic task	
48.	**Decision support system for inventory management with an application**	4
49.	**Methodologies for expert system development**	2
50.	**Analysis of data mining methodologies**	1
51.	**Design of an oblique decision tree classifier using fuzzy measures**	8
52.	**Diagnostic method using non_monotonic reasoning**	1
53.	**Cryptography for computer networks security**	2
54.	**An object_oriented framework for accounting systems**	2
55.	**A tool for knowledge discovery in life insurance data**	2
56.	**Developing a text retrieval system with a natural language interface**	2
57.	**Data mining for improving data bases capabilities**	2
58.	**Diagnostic reasoning method using dynamic model**	1
59.	**Authorization system in object_oriented data bases**	1
60.	**Towards an object_oriented data bases**	1
61.	**Intelligent dynamic scheduling for real time systems**	4
62.	**On fuzzy activity network analysis**	1
63.	**A firewalls based scheme for computer networks security**	2
64.	**Learning from analogical examples**	1
65.	**A tool for local cash flow forecasting based on fuzzy neural network**	2
	Total	194

Table 4-3 shows that an average of 3 Translations would be produced for each input.

A list of the system output is given Appendix B.

4.2.3 Experiment III

The purpose of this experiment is to test the system on new titles to be sure that system output is satisfactory. The evaluation methodology is applied on 9 titles from computer science domain. This translation system gives 4 matched translations, See table 4-4.

Table 4-4 **Matched Translation in Experiment III**

Titles	Machine translation	Original translation
Neural networks in forecasting models : Nile river application	شبكات عصبية في نماذج التنبؤ : تطبيق نهر النيل	الشبكات العصبية في نماذج التوقع : التطبيق علي نهر النيل
A new algorithm for learning in neural networks using piecewise linear function	خوارزمية جديدة للتعلم في الشبكات العصبية باستخدام الدالة الخطية المتقطعة	خوار زم جديد لتعليم الشبكات العصبية التي تستخدم دالة خطية متقطعة
An integrated model for internet security using prevention and detection techniques	نموذج متكامل لتأمين شبكة الإنترنت باستخدام تقنيات الوقاية و الكشف	نموذج متكامل لتأمين شبكة الإنترنت باستخدام نظامي الكشف و الوقاية
Exploiting parallelism in computational linguistics applications	استغلال التوازي في تطبيقات اللغويات الحسابية	استغلال التوازي في تطبيقات اللغويات الحسابية

The remaining 5 titles have problems that arise from the divergences and mismatches between source and target phrases. See table 4-5

Table 4-5 Evaluation result in Experiment III

Titles	Machine translation	Original translation	Problem No.	Score
A data communication network over electric power distribution networks	شبكة اتصالات بيانات علي شبكات توزيع القدرة الكهربية	شبكة للاتصالات الرقمية تعمل علي شبكات توزيع القوي الكهربية	2	8.5
Performance analysis of microwave distributed amplifiers	تحليل أداء المكبرات الموزعة الموجة الميكروئية	تحليل أداء المكبرات الموزعة التي تعمل في حيز الموجات الميكروئية	1	7
A new image compression technique using wavelet and vector quantization	تقنية ضغط صورة جديدة باستخدام المويجة و تكمية المتجهه	طريقة جديدة لضغط بيانات الصور باستخدام المويجات و تكمية المتجهات	4	8.25
Image coding techniques using wavelets and fractals	تقنيات تكويد الصورة باستخدام المويجات و الجزئيات	طرق تكويد الصور باستخدام المويجات و الجزئيات	1	10
A new approach for realizing electronic chaos generators	منهج جديد لتحقيق مولدات الفوضى الإلكترونية	اتجاه جديد لتحقيق مولدات الفوضى الإلكترونية	1	10
Total				43.8

The evaluation results are shown in Table 4-5. This table shows that the 5 noun phrases takes 43.8 total score. The correctness of the 9 titles is about 93% in average

Order of the adjective

Two additional situations that belong to the order of adjective problems mentioned above. In the same manner, switching the adjective for the specific patterns makes the solution. Consequently, two transfer rules were added. They are described as follows:

1. Switching an adjective with the noun that precedes it. This is specified by the following transfer rule.

(13) N1:$1,ADJ:$2,N2:$3
⇨ADJ:$2, N1:$2, N2:$3

The following example shows the improvement realized by the addition of this rule.

Title	Machine translation	Original translation
Performance evaluation of microwave distributed amplifier	تحليل أداء المكبرات الموجة الميكروئية الموزعة	تحليل أداء المكبرات الموزعة التي تعمل في حيز الموجات الميكروئية

2. Moving the adjective that are followed by three nouns after the second noun, then adding the preposition "ل" to the second noun. This is specified by the following transfer rule.

(14) ADJ:$1,N1:$2,N2:$3,N3:$4
⇔ N3:$4,ADJ:$1,Sept'ل', N2:$3,N1:$2

The following example shows the improvement realized by the addition of this rule.

Title	Machine translation	Original translation
A new image compression technique using wavelet and vector quantization	تقنية جديدة لضغط الصورة باستخدام المويجة و تكمية المتجه	طريقة جديدة لضغط بيانات الصور باستخدام المويجات و تكمية المتجهات

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

This thesis has been concentrated on issues in the design and implementation of a transfer-based machine translation system, which translates an English noun phrase into Arabic. We found that the translation of the title of scientific text, e.g. thesis and journal, to be closely applicable to this system. A major design goal of this system is that it can be used as a stand-alone tool and can be very well integrated with a general MT system for English sentence. It consists of three main modules, responsible for analysis, transfer, and generation. The system uses an English-Arabic bilingual dictionary, an English monolingual, and an Arabic monolingual dictionary. The system is implemented in Prolog and the parser is written in DCG formalism.

The development of the parser is a two-step process. In the first step, we acquire the rules that constitute a grammar for the English noun phrase that gives a precise account of what it is for a noun phrase to be grammatical. This was performed by examining large number of titles in the computer science domain. The grammar covers simple and compound noun phrases. A compound noun phrase is two or more simple noun phrase connected by a separator. A separator helps in recognizing the beginning of a new noun phrase. The second step of the analysis component is to implement the parser that assigns grammatical structure onto input noun phrase. In order to implement the parser, it was needed to perform morphological analysis on the inflected Arabic words. An English monolingual dictionary was also needed to successfully implement the morphological analyzer. The morphological analyzer returns to the parser the words in its primitive form with some additional information such as the number of a noun. Entries of the English dictionary can be prepositions, conjunctions, punctuation symbols, adjectives, and nouns. Considering nouns, it is stored in the dictionary in their singular form. In addition, the Arabic dictionary also includes entries for compound forms, usually for technical terms, and irregular nouns.

The transfer module systems rely on mappings between the surface structure of sentences: a collection of tree-to-tree transformations is applied recursively to the analysis tree of the source language in order to construct a target language analysis. The tree-to-tree transformation algorithm is a recursive, non-deterministic, top-down process in which one side of the tree-to-tree transfer rules is matched against the input structure, resulting in the structure on the right-hand-side. Just as the analysis component has a dictionary so also the transfer component has a bilingual dictionary. This dictionary relates the primitive form of the English lexical units to the primitive form of the Arabic lexical units.

In our noun phrase translator, the actual translation occurs in the transfer phase. The problems encountered while designing this phase comes from the divergences and mismatches between source and target sentences. In general, several different transfer rules will be able to apply to a structure, giving alternative (not necessarily correct) translations. Alternative translations may be inevitable because the translation rules produce correct target structure in some cases while they produce non-acceptable-grammatically incorrect- translations in the others.

The generation component is responsible for synthesizing the Arabic words in its right form. The generated syntactic tree is traversed in a depth-first manner to produce the surface Arabic noun phrase.

The input to this module is the origin of the Arabic words, which is passed by transfer phase with some information about each word. The output is the Inflected Arabic word. This module synthesizes right number and definition of a noun, and the right definition and gender of an adjective. This process includes an Arabic monolingual dictionary. Considering nouns and adjectives, they are stored in the dictionary in their singular masculine form.

In order to evaluate the correctness of our MT system we developed an evaluation methodology. This methodology is based on a comparison between the system output with the original translation of the input text. In order to improve the translation output the evaluation methodology is applied on successive stages that includes a cycle of translation, error identification, correction, and re-translation until no more changes can be made cost effectively and without diminishing the quality of other areas of the system.

Three experiments were conducted. The results observed in first experiment were satisfactory. It shows that 88% of the translation were correct. Then, the system output is improved by solving the main problem encountered in this experiment. In the second experiment the system is run again on the same test set. The obtained results show that an average of 3 translations would be produced for each input noun phrase. In the third experiment the system were tested on new titles. It shows that 93% of the translation were correct.

5.2 Future Work

We have several on going activities, all concerned with extending our thesis work to be more powerful and applicable. In what follows, we present some of these activities.

Implement the semantic analysis phase, which concerns representing the meaning of an individual noun phrase. This task is difficult because there is no universally accepted notation has been devised for semantics. This phase is essential for solving semantics ambiguities.

Conduct a comparative study by reimplementing the translation using other translation approaches such as interlingual and example-based translation.

Integrate this system into some other applications such as systems for teaching Arabic as a foreign language.

As noun phrases are frequently used in scientific and technical documents, we would also like to look into the issue of extending our system to be able translate scientific and technical material such as computer manuals.

REFERENCES

Amgarten M. and Merz D. A.: Towards a systematic Evaluation of Machine Translation System, University of Zurich Department of Computer Science, Computational Linguistics Winterthurerstr. 190, CH-8057 Zurich, 1997 mamgar@cl.unizh.ch. merz@ifi.unizh.ch

Apresian J., Boguslavskij I. Iomdin L., Lazurskij A., Sannikov V. and Tsinman L. ETAP-2: The linguistics of a Machine Translation system, META, XXXVII 4 :97-112, 1992

Arnold D.J. and Sadler L.G: The theoretical basis of MiMo. Machine Translation, 5:195-222, 1990

Arnold D.J. "Sur la Conception du Transfert", La Traductique, P. Bouillon and A. Clas, eds., Les Presses de l'universite de Montreal, Montreal, 1993, pp 64-76. A version of this - in English "Transfer"—can be found in http://clwww.essex.ac.uk/~doug/transfer.ps.gz. Working Papers in Language Processing, 40, Dept. of Language and Linguistics, University of Essex, 1992.

Arnold D.J., Lorna Balkan, Siety Meijer, R.Lee Humphreys and Louisa Sadler Machine Translation: an Introductory Guide, Blackwells-NCC, London, 1994, ISBN: 1855542-17x. Available at http://clwww.essex.ac.uk/~doug/book/book.html.

Ball C. Introduction to Computational Linguistics, Lecture 11: Natural Language Processing, 1995 http://www.georgetown.edu/cball/ling361/ling361_nlp1.html

Baker, J.K. Stochastic Modeling for Automatic Speech Understanding. In: Reddy, R. A. Ed. Speech Recognition. Academic Press, New York, NY.1979

Bennett S.W. and Slocum J. METAL: The LRC Machine Translation system. In Slocum J., editor, Machine Translation systems, pp. 111-140. CUP, Cambridge,1988

Bond F., Ogura K. and Ikehara S.: Countability and number in Japanese to English Machine Translation, NTT Communication Science Laboratories 1-2356 Take, Yokosuka, Kanagawa-ken, Japan 238-03, 1995{bond,ogura,ikehara}@nttkb.ntt.jp

Brown P. F., Cocke J., Stephen A. Pietra D., Vincent J. Pietra D., Jelinek F., John D. Lafferty,. Mercer R. L, and Roossin P. S.: A Statistical Approach To Machine Translation 1990.

Brown P.F. et al, Aligning Sentences in Parallel Corpora, in Proceeding Of the 29th Annual Meeting of the Acl, pp 169-176 1991

Brown P.F. et al., The Mathematics of Statistical Machine Translation: Parameter Estimation. Computational Linguistics, pp 263-311 June 1993

Carbonell, J.G., Cullingford R. and Gershman A. Steps Towards Knowledge-based Machine Translation. IEEE Translations on Pattern Analysis and Machine Intelligence, 3 :376-392, 1981.

Cicekli I. and Altay Guvenir H.: Learning Translation Rules From A Bilingual Corpus, 1996.

Doug Arnold, Lorna Balkan, Siety Meijer, R. Lee Humphreys, Louisa Sadler : Machine Translation An Introductory Guide 1994.

Erk H.: Zur Lexik wissenschaftlicher Fachtexte, Verben - Bd 1, Substantive- Bd 2. Munchen: Hueber 1972

Farwell D., Wilks Y. ULTRA: A Multilingual Translator. In S. Nirenburg ed-in- chief , proceedings of Machine Translation Summit III, Washington and Pittsburgh: Center for Machine Translation, Carnegie Mellon University, 19-24. 1991

Ferguson J. D. Hidden Markov Analysis: An Introduction. In: Ferguson J. D. ed. , Hidden Markov Models for Speech. IDA-CRD, Princeton, NJ.1980

Furuse O. and Iida H.: An example-based method for transfer-driven Machine Translation. In 4th International Conference on Theoretical and Methodological Issuse in Machine Translation, pp 139-150, Montreal, Canada 1992

Gale W.A. and Church K.W.: A Program for Aligning Sentences in Bilingual Corpora, in Proceeding Of the 29th Annual Meeting of the ACL., pp 177-184 1991

Garside R.G.: Leech G.N. and Sampson G.R. The Computational Analysis of English: A corpus-based Approach. Longman, NY. 1987.
Hammond K.J.: Ed. Proceedings : Second Case-based Reasoning Workshop. Pensacola Beach, Fl: Morgan Kaufmann, pp 636-646 1989

Hayes P. and Weinstein S. CONSTRUE/TIS : A System for Content-based Indexing of a database of News Stories. In Proceeding of the International Conference on Innovative Applications of AI, Washington, 1-5 1990

Hutchins W.J and Somers H.L : An Introduction to Machine Translation. Academic Press, London, 1992

Hubert Lehmann, Nikolaus Ott : Translation Relation And The Combination Of Analytical And Statistical Methods In Machine Translation .

Ian J. McAllen: Example-Based Machine Translation Using Connectionist Matching.

Isabelle P. Machine Translation at the TAUM group. In M. King, editor, Machine Translation Today: The state of the Art, Proceedings of the Third Lugano Tutorial, 1984, pp 247-277. Edinburgh University Press, Edinburgh, 1987.

Justeson, J. and Katz, S. (1995) 'Technical Terminology: Some Linguistic Properties and an Algorithm for Identification in Text', in Natural Language Engineering, Vol. 1, No 1: 9-27

Kaji H. HICATS/JE: A Japanese-to-English Machine Translation System Based on Semantic. In M. Nagao ed.-in-chief , Machine Translation Summit proceedings of MT Summit I, Hakone, Japan, 1987 . Tokyo: Ohmsha, 101-106 1989

Kaji H., Kida Y. and Morimoto Y. : Learning Translation Templates from Bilingual Text. Proc. Coling., pp 672-678 992

Kay M.: Machine Translation, Stanford University, Palo Alto, CA, 1997
file:///Dl/translation/html/MT/Kay.html

Knight K. And Steve K. Luk : Building a Large-Scale Knowledge Base for Machine Translation .

Kitano H.: A Comprehensive and Practical Model of Memory-based Machine Translation. In Ruzena Bajcsy Ed. Proceedings of the Thirteeth International Joint Conference on Artificial Intelligence, Morgan Kaufmann V. 2 pp 1276-1282 1993

Kolodner J. L.: Ed. Proceedings of a Workshop on Case-Based Reasoning. Clearwater Beach, Fl: Morgan Kaufmann 1988

Landsbergen J. the Rosetta Project. In Proceedings of MT Summit II, Munich, pp 82-87 1989

Lambros Cranias, Harris Papageorgiou, Stelios Piperidis : A Matching Technique In Example-Based Machine Translation 10 Aug 1995 .

Maas H.D. The MT system SUSY. In Margaret King, editor, Machine Translation Today: The state of the Art, Proceedings of the Third Lugano Tutorial, 1984, pp 209-246. Edinburgh University Press, Edinburgh, 1987

McCord M.C. Design of LMT: A Prolog-based Machine Translation system. Computational Linguistics, 15 1: 33-52, 1989

Medin D. L. and Schaffer, M.M.: Context theory of classification learning. Psychological Review, 85 pp 207-238 1988

Meyer I., Onyshkevych B. and Carlson L. Lexicographic Principles and Design for knowledge-based Machine Translation. Technical Report CMU-CMT 90-118, Center for Machine Translation, Carnegie Mellon University, Pittsburgh, and Pa 1990

Muraki K. PIVOT: Two-phase Machine Transltion system. In M. Nagao ed-in-chief , Machine Translation Summti, Tokyo: Ohmsha pp 113-115 1989

Nagao M. A.: Framework of a Mechanical Translation between Japanese and English by Analogy Principle 1984

Nirenburg S., Nirenburg I. and Reynolds J. the Control Structure of POPLAR: a Personality-Oriented Planner. In Proceedings of Third Annual Conference on Intelligent systems, Oakland University, Rochester,Mich. 1985

Nirenburg S. and Goodman K. Treatment of Meaning in MT Systems. In Proceedings of the Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages, University of Texas, Austin, pp 171-188. 1990

Nirenburg S., Pustejovsky J. Processing Aspectual Semantics. In Proceedings of Annual Conference of the Cognitive Science Society, Monteral, pp 658-665 1988

Osamu Furusse And Hitoshi IIDA : An Example-Based Method for Transfer-Driven Machine Translation .

Ram A.: Indexing, Elaboration and Refinement: Incremental Learning of Explanatory Cases. In Janet L. Kolodner ed. , Case-Based Learning, Kluwer Academic Publishers 1993

Sadler V. and Vendelmans R.: Pilot Implementation of a Bilingual Knowledge Bank. Proc. Of Coling, pp 449-451 1990

Sampson G.R. A Stochastic Approach to Parsing. Proceedings of the 11th International Conference on Computational Linguistic. Pp 151-155 1986

Sergei Nirenburg, Jaime Carbonell, Masaru Tomita, Kenneth Goodman: Machine Translation A Knowledge-Based Approach

Sharman R.A., Jelinek F. and Mercer R.L. Generating a Grammar for Statistical training. In: Proceedings of the IBM Conference on Natural Language Processing. Thornwood, NY. 1988

Sharp R. CAT-2- implementing a formalism for multi-lingual MT. In Proceedings of the 2nd International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages, pp 76-87 Carnegie Mellon University, Center for Machine Translation, Pittsburgh, USA, 1988

Sinclair J.M. Lexicograhic Evidence. In: Ilson R. ed. Dictionaries, Lexicography and Language Learning. Pergamon Press, New York, NY. 1985

Slocum J., Bennett W.S., Bear J., Morgan M. and Root R.. Meta: The LRC Machine Translation system. In M. King, editor, Machine Translation Today, pp 319-350. Edinburg University Press, Edinburgh, 1987

Sumita E. and Iida H.: Experiments and Prospects of Example-Based Machine Translation. Proc. Of the 29th Annual Meeting of the Association for Computational Linguistics, pp 185-192 1991

Takeda K.: Pattern-Based Context-free Grammars for Machine Translation Tokyo Research Laboratory, IBM Research 1623-14 Shimostsuruma, Yamato, Kanagawa 242, 1996 takeda@trl.vnet.ibm.com

Tomita M. a Universal Parser Architecture for knowledge-based Machine Translation. In Proceedings of 10th Annual Meeting of the International Joint Conference for Artifical Intelligence IJCAI-87, Milan, pp 718-721 1987

Trujillo A. : Translation Engines Techniques for Machine Translation, Springer (1999)

Uchida H. ATAS-II: A Machine Translation System Using Conceptual Structure as an Interlingual. In M. Nagao, M. rd-in-chief , Machine Translation Summit. Tokyo: Ohmsha proceedings of the 1987 summit, Hakone, Japan , pp 93-100. 1989

Van Noord G., Dorrepaal J., van der Eijk P., Florenza M. and des Tombe L. The MiMo2 research system. In Third International Conference on Theoretical and Methodological Issues in Machine Translation, pp 11-13 June 1990, pp 213-124, Linguistics Research Center, Austin, Texas, 1990

Vauquois B. and Boitet C. Auutomated Translation at Gernoble University. Computational Linguistics 11:28-36 1985

Wilks Y. The Stanford Machine Translation and Understanding Project. In R. Rustin ed. , Natural Language Processing, Courant Computer Science Symposium, No.8 . New York: Algorithmic Press pp 243-290 1973

Witkam T. DLT-Ditributed Language Translation--A Multilingual Facility for Videotex Information Networks. Technical Report, BSO Buro voor Systemontwikkeling , Utrecht, The Netherlands. 1983 .

Witkam T. Interlingual Machine Translation --An Industrial Initiative. In M. Nagao, M. ed.-in-chief , Machine Translation Summit. Tokyo: Ohmsha proceedings of the 1987 summit, Hakone, Japan , pp 128-132. 1989

Wu D.: Grammarless Extraction of Phrasal Translation Examples From Parallel Texts, Proceedings of the Sixth Int. Conf. On Theoretical and Methodological Issues in Machine Translation Leuven Belguim 1995

Yuh S., Jung H., Chang W., Kim T. and Park D.I.: From To/JK: A Japanaese-Korean Machine Translation System, Machine Translation Lab., Systems Engineering Research Institute, 1997 {shyuh,jhm,wchang,twkim,dipark}@seri.re.kr

Arabic References

1- ملخص قواعد اللغة العربية تأليف فؤاد نعمة
مرجع كامل لقواعد النحو و الصرف أ عد بأسلوب شيق و مبتكر وبصورة مبسطة
سهلة و مرتبة مع التوضيح بالأمثلة و الجداول(طبعة مايو 1973)

APPINDEX **A**

THE SYSTEM OUTPUT DURING DEVELOPMENT STAGE

knowledge base verification
تحقيق قاعدة المعرفه

problems in software maintenance
مشكلات في صيانه البرمجيات

design and implementation of a system for program verification
تصميم و تنفيذ نظام لتحقيق البرنامج

analysis and evaluation to some time stamp based algorithms in data base concurrency control
تحليل و تقييم لبعض خوارزميات قائمة علي استخدام الوقت في مراقبة توازى قاعدة البيانات

data modeling in data bases and conversion between models
نمذجه البيانات في قواعد البيانات و تحويل بين النماذج

knowledge base techniques in an arabic syntax analysis environment
تقنيات قاعدة المعرفه في بيئه تحليل بنيه لغه عربية

assessment of structure using neural networks approach
تقييم الإنشاء باستخدام منهج الشبكات العصبية

a natural language interface for querying a relational data base systems
واجهة لغة طبيعية لاستفسار عن نظم قاعدة بيانات علائقية

performance evaluation of routing in integrated service digital network

تقييم أداء المسارات في الشبكة الرقمية الخدمية المتكاملة

computer based corporate financial simulation modeling system
نظام نمذجه المحاكاه المالية الشركية القائمة علي حاسب

decision support system for financial management
نظام دعم القرار للإدارة المالية

an intelligent assistant for extracting knowledge from a data base
مساعد ذكي لاستخلاص المعرفه من قاعدة بيانات

conversion of isolated data bases into a distributed data base system
تحويل قواعد البيانات المنفصلة الي نظام قاعدة بيانات موزعة

providing an intelligent behaviour to relational data base system
اعطاء سلوك ذكي لنظام قاعدة البيانات العلائقية

data recovery system in relational data bases
نظام إستعاده البيانات في قواعد البيانات العلائقية

data integrity control system for relational data bases
نظام مراقبة تكامل البيانات لقواعد البيانات العلائقية

contribution to the solution of fingerprint identification problem
مساهمة لحل مشكلة مطابقة البصمة

expert system techniques for software development of a computer_aided arabic translation system
تقنيات النظام الخبير لتطوير برمجيات نظام ترجمة لغه عربية بمساعدة الحاسب

decision support for clinical knowledge base systems
دعم القرار لنظم قاعدة المعرفه الإكلينكية

neural networks for economic decision making
شبكات عصبية لصنع القرار الإقتصادي

a fuzzy logic system for seismic predictions
نظام منطقي فازي لتنبوءات الزلازل

design of a classifier for fingerprint identification
تصميم مصنف لمطابقة البصمة

dependency networks as a knowledge representation scheme
شبكات اعتمادية كطريقة تمثيل معرفه

study of optimization space in feedforward neural networks
دراسة فضاء الأمثليه في الشبكات العصبية الأمامية التغذيه

building an expert system for determining the value of custom purposes
بناء نظام خبير لقيمه تحديد الأغراض الجمركية

sharable and reusable infraknowledge for large scale knowledge base systems
مشاركة و اعادة أستخدام المعرفه الأساسيه لنظم قاعدة معرفه النطاق الواسع

knowledge representation for multitype inference in multistrategy task adaptive learning
تمثيل المعرفه لاستنتاج الأنواع المتعددة في التعلم المقوم مهمة الاسرتيجيات المتعددة

a new language for graphical user interface application
لغة جديدة لتطبيق واجهة المستخدم الرسومي

a comparative study for logic programming models
دراسة مقارنة لنماذج البرمجه المنطقية

integrating constraint logic programming with decision support systems
تكامل البرمجه المنطقية المقيدة مع نظم دعم القرار

routing algorithm for multimedia information transmission over a widearea network
خوارزمية المسارات لنقل معلومات الوسائط المتعددة علي شبكة متسعة

switching techniques for multimedia information transmission over a widearea network
تقنيات التحويل لنقل معلومات الوسائط المتعددة علي شبكة متسعة

performance evaluation of routing algorithms in high speed networks
تقييم أداء خوارزميات المسارات في الشبكات السريعة العالية

synchronization problems facing multimedia transmission
مشكلات التزامن التي تواجه نقل الوسائط المتعددة

deadlock and congestion avoidance in widearea networks
توقف تام و تجنب الإزدحام في الشبكات المتسعة

selection and optimization of appropriate potable water supply for small communities using simulation techniques
اختيار و أمثليه مصدر المياه الصالح للشرب المناسب لمجتمعات صغيرة بإستخدام تقنيات المحاكاه

a comprehensive study on blood banks inventory management system
دراسة شاملة علي نظام إدارة مخزون بنوك الدم

on the recent approaches for solving multiobjective linear programming problems
من المناهج الحديثة لحل مشكلات البرمجه الخطية المتعددة الأهداف

study on multiobjective nonlinear programming problems
دراسة علي مشكلات البرمجه الغير خطية المتعددة الأهداف

decision support system for shortest path algorithms with an application
نظام دعم القرار لخوارزميات المسار الأقصر مع تطبيق

optimizing the net_present_value for cash flows in resource constrained projects
أمثلية صافي القيمه الحاليه لتدفقات نقدية في المشروعات المحدودة المورد

an intelligent multiobjective programming approach for solving rural wastewater management problems
منهج برمجه متعددة أهداف ذكية لحل مشكلات إدارة الصرف الصحى القروي

usage of expert systems in scheduling
استخدام النظم الخبيرة في الجدوله

an expert system for an interactive multiobjective programming
نظام خبير لبرمجه متعددة أهداف تفاعلية

integrating modern information technology into school teaching
تكامل تقنية المعلومات الحديثة الي التعليم المدرسي

intelligent search system for bibliographic services
نظام البحث الذكي للخدمات البيبلوجرافية

integrating expert systems with multimedia
تكامل النظم الخبيرة مع وسائط متعددة

storage and query processing in multimedia data bases systems
تخزين و معالجة الاستفسار في نظم قواعد بيانات الوسائط المتعددة

APPINDEX **B**

THE SYSTEM OUTPUT DURING EVALUATION STAGE

B-1 The Output From Noun Phrases Used In Experiment I

distributed data bases system: features and impact on applications
نظام موزع لقواعد البيانات: خصائص و تأثير علي التطبيقات
نظام موزع لقواعد البيانات: خصائص و تأثير للتطبيقات
نظام قواعد البيانات الموزع: خصائص و تأثير علي التطبيقات
نظام قواعد البيانات الموزع: خصائص و تأثير للتطبيقات
نظام قواعد البيانات الموزعة: خصائص و تأثير علي التطبيقات
نظام قواعد البيانات الموزعة: خصائص و تأثير للتطبيقات

an authoring shell for intelligent tutoring systems
أداة تأليف في نظم التعليم الذكي
أداة تأليف في نظم التعليم الذكية
أداة تأليف لنظم التعليم الذكي
أداة تأليف لنظم التعليم الذكية

automatic generation of explanation for expert systems
توليد أتوماتيكي للتوضيح في النظم الخبيرة
توليد أتوماتيكي للتوضيح للنظم الخبيرة

a concurrent constraint_logic_programming language
لغة برمجة منطقية مقيدة متزامنة

behaviour simulation of protein based digital systems
محاكاه سلوك النظم الرقمية القائمة علي البروتين

a computer based clinical information system
نظام معلومات إكلينكي قائم علي حاسب
نظام معلومات إكلينكية قائمة علي حاسب

neural networks approach for controlling the operation of power transformer differential relay

منهج الشبكات العصبي في تحكم في تشغيل المتمم الفرقي لمحول القدرة
منهج الشبكات العصبي في تحكم في تشغيل المتمم الفرقي محول القدرة
منهج الشبكات العصبي في تحكم في تشغيل متمم المحول الفرقي القدرة
منهج الشبكات العصبي لتحكم في تشغيل المتمم الفرقي لمحول القدرة
منهج الشبكات العصبي لتحكم في تشغيل المتمم الفرقي محول القدرة
منهج الشبكات العصبي لتحكم في تشغيل متمم المحول الفرقي القدرة
منهج الشبكات العصبية في تحكم في تشغيل المتمم الفرقي لمحول القدرة
منهج الشبكات العصبية في تحكم في تشغيل المتمم الفرقي محول القدرة
منهج الشبكات العصبية في تحكم في تشغيل متمم المحول الفرقي القدرة
منهج الشبكات العصبية لتحكم في تشغيل المتمم الفرقي لمحول القدرة
منهج الشبكات العصبية لتحكم في تشغيل المتمم الفرقي محول القدرة
منهج الشبكات العصبية لتحكم في تشغيل متمم المحول الفرقي القدرة

expert system development tool based on routine design generic task

أداة خبيرة لتطوير النظام القائم علي المهمة العامة للتصميم الروتيني
أداة خبيرة لتطوير النظام قائم علي المهمة العامة للتصميم الروتيني
أداة تطوير النظام الخبير القائم علي المهمة العامة للتصميم الروتيني
أداة تطوير النظام الخبيرة القائمة علي المهمة العامة للتصميم الروتيني

diagnostic method using non_monotonic reasoning

طريقة تشخيصية بإستخدام السببية الغير رتيبة

investigating a new constraint solving approach within constraint_logic_programming

بحث منهج جديد لحل القيد في البرمجة المنطقية المقيدة
بحث منهج جديد لحل القيد من خلال البرمجة المنطقية المقيدة
بحث منهج حل قيد جديد في البرمجة المنطقية المقيدة
بحث منهج حل قيد جديد من خلال البرمجة المنطقية المقيدة

cryptography for computer networks security

تشفير في تأمين شبكات الحاسب
تشفير لتأمين شبكات الحاسب

a firewalls based scheme for computer networks security

طريقة قائمة علي حوائط نارية في تأمين شبكات الحاسب
طريقة قائمة علي حوائط نارية لتأمين شبكات الحاسب

performance evaluation of proposed routing techniques for the internet

تقييم مقترح مسارات لأداء التقنيات في شبكة الأنترنت
تقييم مقترح مسارات لأداء التقنيات لشبكة الأنترنت
تقييم أداء تقنيات المقترح للمسارات في شبكة الأنترنت
تقييم أداء تقنيات المقترح للمسارات لشبكة الأنترنت
تقييم أداء تقنيات مقترحة للمسارات في شبكة الأنترنت
تقييم أداء تقنيات مقترحة للمسارات لشبكة الأنترنت
تقييم أداء التقنيات المقترح المسارات في شبكة الأنترنت
تقييم أداء التقنيات المقترح المسارات لشبكة الأنترنت
تقييم أداء التقنيات المقترحة المسارات في شبكة الأنترنت
تقييم أداء التقنيات المقترحة المسارات لشبكة الأنترنت
تقييم أداء تقنيات المسارات المقترح في شبكة الأنترنت

تقييم أداء تقنيات المسارات المقترح لشبكة الأنترنت
تقييم أداء تقنيات المسارات المقترحة في شبكة الأنترنت
تقييم أداء تقنيات المسارات المقترحة لشبكة الأنترنت
تقييم الأداء لتقنيات مسارات مقترحة في شبكة الأنترنت
تقييم الأداء لتقنيات مسارات مقترحة لشبكة الأنترنت

a study on recent approaches in goal_programming with an application
دراسة علي مناهج حديثة في البرمجة الهدفية بتطبيق
دراسة علي مناهج حديثة في البرمجة الهدفية مع تطبيق
دراسة علي مناهج حديثة في برمجة الاهداف بتطبيق
دراسة علي مناهج حديثة في برمجة الاهداف مع تطبيق
دراسة لمناهج حديثة في البرمجة الهدفية بتطبيق
دراسة لمناهج حديثة في البرمجة الهدفية مع تطبيق
دراسة لمناهج حديثة في برمجة الاهداف بتطبيق
دراسة لمناهج حديثة في برمجة الاهداف مع تطبيق

a tool for local cash flow forecasting based on fuzzy neural networks
أداة في تنبؤ التدفق النقدي المحلي القائم علي الشبكات العصبية المشوشة
أداة لتنبؤ التدفق النقدي المحلي القائم علي الشبكات العصبية المشوشة

project resource allocation with the support of an expert system based on fuzzy logic
تحديد مورد المشروع بدعم نظام خبير قائم علي المنطق المشوش
تحديد مورد المشروع مع دعم نظام خبير قائم علي المنطق المشوش

an object_oriented framework for accounting systems
اطار شيئي التوجه في نظم المحاسبية
اطار شيئي التوجه لنظم المحاسبية

a tool for knowledge discovery in life insurance data
أداة في اكتشاف المعرفه في بيانات تأمين الحياة
أداة لاكتشاف المعرفه في بيانات تأمين الحياة

developing a text retrieval system with a natural language interface
تطوير نظام استرجاع نص بواجهة لغة طبيعية
تطوير نظام استرجاع نص مع واجهة لغة طبيعية

data_mining for improving data bases capabilities
تنقيب عن البيانات في تحسين قدرات قواعد البيانات
تنقيب عن البيانات لتحسين قدرات قواعد البيانات

an object_oriented approach for building hypermedia systems
منهج شيئي التوجه في بناء نظم الوسائط المترابطة
منهج شيئي التوجه لبناء نظم الوسائط المترابطة

visual composition for intelligent tutoring systems
تجمع مرئى في نظم التعليم الذكي
تجمع مرئى في نظم التعليم الذكية
تجمع مرئى لنظم التعليم الذكي

تجمع مرئى لنظم التعليم الذكية

enhancing authoring systems with the knowledge base concepts
تحسين نظم التأليف بمفاهيم قاعدة المعرفه
تحسين نظم التأليف مع مفاهيم قاعدة المعرفه

automating and integrating different knowledge acquisition techniques
أتمتة و تكامل تقنيات اكتساب المعرفه المختلف
أتمتة و تكامل تقنيات اكتساب المعرفه المختلفة

diagnostic reasoning method using dynamic model
طريقة السببية التشخيصية بإستخدام النموذج الديناميكي

methodologies for expert system development
مناهج في تطوير النظام الخبير
مناهج لتطوير النظام الخبير

integrating artificial intelligence with project management scheme
تكامل الذكاء الاصطناعى بطريقة إدارة المشروع
تكامل الذكاء الاصطناعى مع طريقة إدارة المشروع

an integer goal_programming model for improving athletics training schedule
نموذج برمجة الاهداف الصحيحة في تحسين جدول التدريب الرياضي
نموذج برمجة الاهداف الصحيحة لتحسين جدول التدريب الرياضي
نموذج برمجة الاهداف صحيح في تحسين جدول التدريب الرياضي
نموذج برمجة الاهداف صحيح لتحسين جدول التدريب الرياضي
نموذج برمجة هدفية صحيح في تحسين جدول التدريب الرياضي
نموذج برمجة هدفية صحيح لتحسين جدول التدريب الرياضي
نموذج برمجة هدفية صحيحة في تحسين جدول التدريب الرياضي
نموذج برمجة هدفية صحيحة لتحسين جدول التدريب الرياضي

analysis of data_mining methodologies
تحليل مناهج التنقيب عن البيانات

authorization system in object_oriented data bases
نظام الصلاحية في قواعد البيانات الشيئية التوجه

automatic knowledge acquisition tool for scheduling systems
أداة أتوماتيكية لاكتساب المعرفه في نظم الجدوله
أداة أتوماتيكية لاكتساب المعرفه لنظم الجدوله
أداة اكتساب المعرفه الأتوماتيكية في نظم الجدوله
أداة اكتساب المعرفه الأتوماتيكية لنظم الجدوله

design of an oblique decision tree classifier using fuzzy measures
تصميم مائل قرار لمصنف الشجرة بإستخدام القياسات المشوشة
تصميم مصنف شجرة مائل للقرار بإستخدام القياسات المشوشة
تصميم مصنف شجرة مائلة للقرار بإستخدام القياسات المشوشة

تصميم مصنف لشجرة قرار مائل بإستخدام القياسات المشوشة
تصميم مصنف لشجرة قرار مائلة بإستخدام القياسات المشوشة
تصميم مصنف شجرة مائل قرار بإستخدام قياسات مشوشة
تصميم مصنف شجرة مائلة قرار بإستخدام قياسات مشوشة
تصميم مصنف شجرة قرار مائل بإستخدام القياسات المشوشة

a knowledge based system for software maintenance
نظام قائم علي معرفه في صيانه البرمجيات
نظام قائم علي معرفه لصيانه البرمجيات

learning from analogical examples
تعلم من الأمثلة الشبيهة

arabic text retrieval and classification using neural network
استرجاع و تصنيف نصوص عربية بإستخدام الشبكة العصبية

a computer_assisted tool for evaluation of pascal programs
أداة بمساعدة حاسب في تقييم برامج الباسكال
أداة بمساعدة حاسب لتقييم برامج الباسكال

an application of combined logic and object_oriented programming paradigms
تطبيق منطق اندماج و نماذج البرمجه الشيئي التوجه
تطبيق منطق اندماج و نماذج البرمجه الشيئية التوجه

call admission control in high speed computer networks
مراقبة أجراء الاتصال في شبكات الحاسب السريع العالي
مراقبة أجراء الاتصال في شبكات الحاسب السريعة العالية

evaluation and improvement of work efficiency in wastewater stations using simulation technique
تقييم و تطوير كفاءة العمل في محطات الصرف الصحى بإستخدام تقنية المحاكاه

computer_simulation for a hospital resources planning
محاكاة باستخدام الحاسب في تخطيط موارد مستشفي
محاكاة باستخدام الحاسب لتخطيط موارد مستشفي

on fuzzy multiobjective transportation problem
من مشكلة نقل المتعددة الأهداف المشوشة

on fuzzy continuous static games
من مباريات ساكنة متصلة مشوشة

decision support system for fuzzy multiple projects resources scheduling
نظام دعم القرار في جدوله موارد المشروعات المتعددة المشوشة
نظام دعم القرار لجدوله موارد المشروعات المتعددة المشوشة

developing a cost function in a competitive market using an operation_research technique

تطوير دالة تكلفة في سوق تنافسي بإستخدام تقنية بحوث عمليات

an interactive fuzzy goal_programming model for multiobjective decisions support systems

نموذج برمجة الاهداف المشوشة التفاعلية في نظم دعم قرارات المتعددة الأهداف
نموذج برمجة الاهداف المشوشة التفاعلية لنظم دعم قرارات المتعددة الأهداف
نموذج برمجة الاهداف مشوش تفاعلي في نظم دعم قرارات المتعددة الأهداف
نموذج برمجة الاهداف مشوش تفاعلي لنظم دعم قرارات المتعددة الأهداف
نموذج برمجة هدفية مشوش تفاعلي في نظم دعم قرارات المتعددة الأهداف
نموذج برمجة هدفية مشوش تفاعلي لنظم دعم قرارات المتعددة الأهداف
نموذج برمجة هدفية مشوشة تفاعلية في نظم دعم قرارات المتعددة الأهداف
نموذج برمجة هدفية مشوشة تفاعلية لنظم دعم قرارات المتعددة الأهداف

building a decision support system for organizations in competitive market

بناء نظام دعم قرار في منظمات في السوق التنافسي
بناء نظام دعم قرار لمنظمات في السوق التنافسي

quality assurance for information systems development

تأكيد الجودة في تطوير نظم المعلومات
تأكيد الجودة لتطوير نظم المعلومات

towards an object_oriented data base

نحو قاعدة بيانات شيئية التوجه

a computer tool for knowledge acquisition in different domains

أداة حاسب في اكتساب المعرفه في المجالات المختلفة
أداة حاسب لاكتساب المعرفه في المجالات المختلفة

knowledge modeling and representation using logic programming and constraints based approach

نمذجه و تمثيل معرفه بإستخدام البرمجه المنطقية و منهج قائم علي القيود

evaluation of flow control in integrated service digital network

تقييم مراقبة التدفق في الشبكة الرقمية للخدمة المتكاملة

a computer_simulation model for analyzing the accounting systems of industrial companies

نموذج محاكاة باستخدام الحاسب في تحليل نظم محاسبية للمنشآت الصناعي
نموذج محاكاة باستخدام الحاسب في تحليل نظم محاسبية للمنشآت الصناعية
نموذج محاكاة باستخدام الحاسب لتحليل نظم محاسبية للمنشآت الصناعي
نموذج محاكاة باستخدام الحاسب لتحليل نظم محاسبية للمنشآت الصناعية

comprehensive study on neural networks applications in mathematical programming

دراسة شاملة علي تطبيقات الشبكات العصبية في البرمجه الرياضية
دراسة شاملة لتطبيقات الشبكات العصبية في البرمجه الرياضية

computer based system for item banks: analysis, design and implementation
نظام قائم علي حاسب في بنوك الاسئلة: تحليل و تصميم و تنفيذ
نظام قائم علي حاسب لبنوك الاسئلة: تحليل و تصميم و تنفيذ

integrating expert systems with multimedia
تكامل النظم الخبير بالوسائط المتعددة
تكامل النظم الخبير مع الوسائط المتعددة
تكامل النظم الخبيرة بالوسائط المتعددة
تكامل النظم الخبيرة مع الوسائط المتعددة

assessment and applicability of software maintenance methodologies
تقييم و امكانية تطبيق مناهج صيانه البرمجيات

developing a tool for the fusion method in object_oriented programming
تطوير أداة في طريقة الدمج في البرمجه الشيئية التوجه
تطوير أداة لطريقة الدمج في البرمجه الشيئية التوجه

an expert system for engineering_insurance
نظام خبير في التأمين الهندسي
نظام خبير للتأمين الهندسي

expert system development tool based on hierarchical classification generic task
أداة خبيرة لتطوير النظام القائم علي المهمة العامة للتصنيف الهرمي
أداة خبيرة لتطوير النظام قائم علي المهمة العامة للتصنيف الهرمي
أداة تطوير النظام الخبير القائم علي المهمة العامة للتصنيف الهرمي
أداة تطوير النظام الخبيرة القائمة علي المهمة العامة للتصنيف الهرمي

a hybrid framework for optimal system design
اطار هجين في تصميم النظام الأمثل
اطار هجين لتصميم النظام الأمثل

agent approach for knowledge base systems
منهج العميل في نظم قاعدة المعرفه
منهج العميل لنظم قاعدة المعرفه

towards a knowledge modeling approach in risk analysis domain
نحو منهج لنمذجه المعرفه في مجال تحليل المخاطر
نحو منهج لنمذجه المعرفه في مجال تحليل المخاطرة

a simulation model for transient storage stations
نموذج محاكاه في محطات التخزين المؤقت
نموذج محاكاه في محطات التخزين المؤقتة
نموذج محاكاه لمحطات التخزين المؤقت
نموذج محاكاه لمحطات التخزين المؤقتة

intelligent dynamic scheduling for real time systems

جدوله ديناميكية ذكي في نظم الوقت الواقعية
جدوله ديناميكية ذكي لنظم الوقت الواقعية
جدوله ديناميكية ذكية في نظم الوقت الواقعي
جدوله ديناميكية ذكية لنظم الوقت الواقعي

on fuzzy activity network analysis
من تحليل شبكة النشاط المشوش

decision support system for inventory management with an application
نظام دعم القرار في إدارة المخزون بتطبيق
نظام دعم القرار في إدارة المخزون مع تطبيق
نظام دعم القرار لإدارة المخزون بتطبيق
نظام دعم القرار لإدارة المخزون مع تطبيق

tools for developing multimedia educational systems
أدوات في تطوير النظم التعليمي الوسائط المتعددة
أدوات في تطوير النظم التعليمية الوسائط المتعددة
أدوات لتطوير النظم التعليمي الوسائط المتعددة
أدوات لتطوير النظم التعليمية الوسائط المتعددة
أدوات في تطوير النظم التعليمية للوسائط المتعددة
أدوات لتطوير النظم التعليمي للوسائط المتعددة

multimedia based educational software
برمجيات تعليمية قائمة علي الوسائط المتعددة

goal_programming within the frame of constraint_logic_programming
برمجة الاهداف في اطار البرمجة المنطقية المقيدة
برمجة الاهداف من خلال اطار البرمجة المنطقية المقيدة
برمجة هدفية في اطار البرمجة المنطقية المقيدة
برمجة هدفية من خلال اطار البرمجة المنطقية المقيدة

B-2 The Output From Noun Phrases Used In Experiment III

a data communication network over electric power distribution networks
شبكة اتصالات بيانات علي شبكات توزيع القدرة الكهربية

neural networks in forecasting models: nile river application
شبكات عصبية في نماذج التنبؤ : تطبيق نهر النيل

a new algorithm for learning in neural networks using piecewise linear function
خوارزمية جديدة في التعلم في الشبكات العصبية بإستخدام الدالة الخطية المتقطعة
خوارزمية جديدة للتعلم في الشبكات العصبية بإستخدام الدالة الخطية المتقطعة

an integrated model for internet security using prevention and detection techniques
نموذج متكامل في تأمين شبكة الأنترنت بإستخدام الوقاية و تقنيات الكشف
نموذج متكامل في تأمين شبكة الأنترنت بإستخدام تقنيات الوقاية و الكشف
نموذج متكامل لتأمين شبكة الأنترنت بإستخدام الوقاية و تقنيات الكشف
نموذج متكامل لتأمين شبكة الأنترنت بإستخدام تقنيات الوقاية و الكشف

exploiting parallelism in computational linguistics applications
استغلال التوازي في تطبيقات اللغويات الحسابية

performance analysis of microwave distributed amplifiers
تحليل أداء المكبرات الموزع الموجة الميكروئية
تحليل أداء المكبرات الموزعة الموجة الميكروئية
تحليل أداء مكبرات الموجة الميكروئية الموزع
تحليل أداء مكبرات الموجة الميكروئية الموزعة

a new approach for realizing electronic chaos generators
منهج جديد في تحقيق مولدات الفوضي الالكتروني
منهج جديد في تحقيق مولدات الفوضي الالكترونية
منهج جديد لتحقيق مولدات الفوضي الالكتروني
منهج جديد لتحقيق مولدات الفوضي الالكترونية

a new image compression technique using wavelet and vector quantization
تقنية جديدة لضغط الصورة بإستخدام المويجة و تكمية المتجه
تقنية جديدة لضغط الصورة بإستخدام تكمية المويجة و المتجه
تقنية ضغط صورة جديدة بإستخدام المويجة و تكمية المتجه
تقنية ضغط صورة جديدة بإستخدام تكمية المويجة و المتجه

image coding techniques using wavelets and fractals
تقنيات تكويد الصورة بإستخدام المويجات و الجزئيات

ملخص الرسالة

في العالم المعاصر هناك حاجة متزايدة لترجمة اللغات الطبيعية. هذا وترجع محاولات ترجمة هذه اللغات إلى نشأة الحاسب. وتعتبر الترجمة الآلية هي محاولة لأتمتة كل أو بعض مراحل معالجة الترجمة من لغة لأخرى.

هذا البحث يعرض محاولتنا لتطوير نظام ترجمة آلية قائمة على النقل للتراكيب الاسمية من اللغة الإنجليزية للعربية. وتعتبر التراكيب الاسمية شائعة الاستخدام في الوثائق التقنية والعلمية. وطبقا لمنهج النقل فان النظام يتكون من ثلاثة مكونات مسئولة عن التحليل والنقل والتوليد. ويقوم مكون التحليل بتمثيل هيكل للتركيب الأسمى طبقا لقواعد اللغة الإنجليزية والمعجم. ويقوم مكون النقل ببناء هيكل للتركيب الأسمى المناظر باللغة العربية وطبقا لقواعد المقارنة. وهذه تتطلب معجم إنجليزي-عربي. ويقوم مكون التوليد ببناء الترجمة النهائية. وهذه تتطلب قواعد توليد التركيب الأسمى العربي بالإضافة إلى معجم عربي.

ولعل أحد الأهداف الرئيسية لتصميم هذا النظام هو إمكانية استخدامه كأداة قائمة بذاتها بالإضافة الى إمكانية تكاملة مع نظام أعم لترجمة اللغة الإنجليزية الى العربية. وقد تم تنفيذ هذا النظام بلغة البرلوج كما تم بناء محلل التركيب الأسمى باستخدام نموذج د.س.ج والذي يتم ترجمته مباشرة إلى لغة البرلوج. وقد تم اختبار لهذا النظام على تراكيب اسمية من الواقع. هذا وتتضمن الرسالة على خبرتنا في تطوير نظام ترجمة آلية وكذلك على عرض لنتائج تطبيقها على عناوين رسائل ومجلات علمية

الترجمة الآلية للتراكيب الاسمية: من الإنجليزية إلي العربية

إعـــــــــــــــــــــــداد

عزة عبد المنعم محمد / المهندسة

رسالة مقدمة إلى كلية الهندسة، جامعة القاهرة
كجزء من متطلبات الحصول على درجة الماجستير
في
هندسة الحاسبات (طبقا للائحة)

كلية الهندسة ، جامعة القاهرة
الجيزة ، جمهورية مصر العربية
2000

الترجمة الآلية للتراكيب الاسمية: من الإنجليزية إلي العربية

إعـــــــــــــــــداد

المهندسة / عزة عبد المنعم محمد

رسالة مقدمة إلى كلية الهندسة، جامعة القاهرة
كجزء من متطلبات الحصول على درجة الماجستير
في
هندسة الحاسبات

تحت إشـــراف

ا.د. أحـمد عبد الواحد رافع
كلية الحاسبات والمعلومات
قسم علوم الحاسب
جامعة القاهرة

ا.د.م هـدى بركة
كلية الهندسـة
قسم هندسة الحاسبات
جامعة القاهرة

كلية الهندسة ، جامعة القاهرة
الجيزة ، جمهورية مصر العربية
2000

الترجمة الآلية للتراكيب الاسمية: من الإنجليزية إلي العربية

إعداد
عزة عبد المنعم محمد

رسالة مقدمة إلى كلية الهندسة، جامعة القاهرة
كجزء من متطلبات الحصول على درجة الماجستير
في
هندسة الحاسبات

يعتمد من لجنة الممتحنين :

الأستاذ الدكتور : هدى أنيس بركة

الأستاذ الدكتور : أحمد عبد الواحد رافع

الأستاذ الدكتور : على حسن فهمي

الأستاذ الدكتور : أيمن الدسوقى إبراهيم

كلية الهندسة ، جامعة القاهرة
الجيزة ، جمهورية مصر العربية
2000