Syntax Analysis and Machine Translation of Bangla Sentences

Md. Musfique Anwar, Mohammad Zabed Anwar and Md. Al-Amin Bhuiyan
Dept. of Computer Science & Engineering, Jahangirnagar University, Bangladesh

Abstract
This paper addresses a method to analyze syntactically Bangla sentence using context-sensitive grammar rules which accept almost all types of Bangla sentences including simple, complex and compound sentences and then interpret the input Bangla sentence to English using the NLP conversion unit. The grammar rules employed for this system allow parsing five categories of sentences according to Bangla intonation. The system is based on analyzing an input sentence and converting into a structural representation (SR). Once the SR is created for a particular sentence it is then converted to corresponding English sentence by NLP conversion unit. For conversion, the NLP conversion unit takes help of the Corpus. The effectiveness of this method has been justified over the demonstration of different Bangla sentences with 28 decomposition rules and the success rates in all cases are over 90%.

Keywords:
Context-sensitive grammar, NLP conversion, Structural representation (SR), Corpus, etc.

1. Introduction

The goal of Bangla machine translation is to establish a system that will analyze, understand, and generate languages which humans use naturally, so that eventually it will be able to address the computer as if it were addressing another person. Since text can contain information at many different granularities, from simple word or token-based representations, to rich hierarchical syntactic representations, to high-level logical representations across document collections, this paper seeks to work at the right level of analysis for the application concerned.

Generally machine translation scheme consists of two major phases. These are analysis and generation. In analysis phase, each word of a sentence is checked lexically and then classified into syntactic categories such as noun, pronoun, adjective, verb etc. These words of syntactic categories can be combined with other words to form syntactically coherent groupings or constituents. A set of constituents which behave the same or share the same functions and distribution is normally called phrasal categories. Actually each sentence is composed of one or more phrases. So if we can identify the syntactic constituents of sentences, it will be easier for us to obtain the structural representation of the sentence [1, 2].

A fair amount of grammar rules have been proposed earlier for Bangla sentence analysis, which starts parsing by breaking a sentence into noun phrase and verb phrase and suggested the way to develop an MT dictionary, which can be a useful tool in the lexical phase analysis [1]. Many works have done on syntax analysis of different types of Bangla sentences with the help of a transformational generative grammar. Those works also showed ways to parse different types of Bangla sentences and proposed an algorithm for the parser [2]. M.M. Hoque and M. Ali developed a parsing methodology for Bangla natural language sentences [1]. They show how phrase structure rules can be implemented by top-down and bottom-up parsing approach using new types of grammars called context-free grammars (CFGs) to parse simple sentences of Bangla. A comprehensive approach for Bangla syntax analysis was proposed by L. Mehedy, S.M. Arifin and M. Kaykobad. They have proposed some context-sensitive grammar rule to parse all types of sentences, including complex, compound, exclamatory and optative sentences. Bangla grammar has an inherent property in forming the verbs, that is, unlike the English grammar, various necessary information of a sentence such as the tense, the person, the mode of verb (wµqvi fve) etc. can be extracted from a finite verb. In this project, the parser accept all Bangla sentences that are syntactically correct, parse each of them using the rules of the proposed grammar, or report an error otherwise. This paper focuses on the formation and use of the grammar rules to be used by the parser in the syntax analysis phase [1].

This paper implements a technique to parse Bangla sentences in a new approach using context-sensitive grammar rules that accept almost all types of Bangla sentences. The most common way to represent grammar is as a set of production rules which states how the parts of speech can put together to make grammatical, or "well-formed" sentences. To identify the grammar of a sentence, individual attributes of words are retrieved and checked with grammar rules of Bangla. If they match then the attributes of the words are combined to form the grammatical features of the sentence. The output of Tokenizer is matched with grammatical rules of Bangla. If a rule whose right-hand side matches with a token, the token is assigned with the appropriate category (parts of speech).

Manuscript received August 5, 2009
Manuscript revised August 20, 2009

2. Fundamentals of Language

A formal language is defined as a set of strings. Each string is a concatenation of terminal symbols [3]. Formal languages, such as C and Java, have strict mathematical definitions. In contrast to formal languages, the natural languages such as Bangla, English, Japanese, have no strict definition but are used by a community of speakers [3].

2.1 Phrase Structure Grammar

Most grammar rule formalisms are based on the idea of phrase structure – that strings are composed of substrings called phrases, which come in different categories. A sentence is a group of words, which makes a complete sense. To express a complete sense, a sentence must have a subject phrase and a predicate phrase. The subject is the part, which names the person or thing we are speaking about. And the predicate is the part, which tells something about the subject [3].
There are three types of phrases in Bangla- Noun phrase, Adjective Phrase and Verb Phrase. Simple sentences are composed of these phrases. Complex and compound sentences are composed of simple sentences [4].

2.2 Analysis

Sentence can be analyzed into three main parts: (i) Syntactic interpretation (or Parsing), (ii) Semantic interpretation and (iii) Pragmatic interpretation.
Parsing is the process of building a parse tree for an input string [3]. The interior nodes of the parse tree represent phrases and the leaf nodes represent words. Semantic interpretation is the process of extracting the meaning of a sentence as an expression in some language representation. In this analysis phase, certain checks are made to ensure that the discrete input components fit together meaningfully. Pragmatic interpretation takes into account the fact that the same words can have different meaning in different situations [2] [3].

2.3 Grammar

A grammar is a finite set of rules that specifies a language [3]. Natural languages have no official grammar, but linguists strive to discover properties of the language by a process of scientific inquiry and then to codify their discoveries in a grammar. After creating Lexicon, the words with different phrases in the grammar are combined. We may use some non-terminal symbols to define different kinds of phrases such as sentence (S), prepositional phrase (PP), noun phrases (NP) etc. A set of non-terminal symbols categorize sub-phrases of the language. For example, a rewrite rule is shown bellow [3]:
Sentence → NounPhrase | VerbPhrase

3. Bangla Grammatical Information Review

This section discusses some essential grammatical issues about different parts of speech of Bangla language [5]. Though any English word may fall in one of eight categories: noun, pronoun, adjective, adverb, verb, preposition, conjunction and interjection, a Bangla word may be in one of five categories: noun, pronoun, adjective, verb and indeclinable. Here, adverb is considered as adjective and the type indeclinable is concerned as preposition, conjunction and interjection [6].

Nouns
Bangla and English nouns may be concrete or abstract. Concrete nouns can be classified as proper noun (Bangladesh — evsjv‡`k), common noun (Student — QvÎ), material noun (Water — cvwb) and collective noun (Team — `j). In Bangla, abstract noun can be classified as proper abstract noun (forgiveness —¶gv) and verbal noun-the name of the works (stretching — cÖmvwiZ K‡i).

Pronouns
There are ten different types of pronouns in Bangla language. They are: (i) Personal (I — Avwg), (ii) Reflexive (Myself — ¯^qs), (iii) Near indicating demonstrative (This — GB), (iv) Far indicating demonstrative (That — H), (v) Collective (All — mKj), (vi) Interrogative (What — wK), (vii) Indefinite (Some — wKQy), (viii) Reciprocal (Each other — ci¯úi), (ix) Relative (Who — whwb) and (x) Others indicating (Other — Ab¨).

Adjectives
Adjectives fall into four subcategories: proper adjective (Rvcvbx †iveU), adjective of quality (g‡bvig cwi‡ek), adjective of quantity (wØ¸b), pronominal adjective (†h †Kvb m`m¨ Dcw¯'Z _vK‡jB n‡e). There are four types of adverbial adjectives in Bangla language. They are: adjective that modifies a verb (`ªyZ MvI), adjective that modifies another adjective (AZ¨š PgrKvi cwi‡ek), adjective that modifies an indeclinable (†m Avgvi gZ ax‡i nv‡U), adjective that modifies a sentence (cwikªgx †jv‡KivB mwZ¨Kvi Ávbx).

Verbs
Verbs may be classified in six categories. They are: (i) Intransitive verb (wk¶v_©xiv co‡Q), (ii) Transitive verb (wk¶v_x©iv eB co‡Q), (iii) Di-transitive verb (wk¶K †Q‡jwU‡K GKwU eB Dcnvi w`‡q‡Q), (iv) Causative verb (wkíx f³‡`i Mvb †kvbvb), (v) Compound verb (a nonfinite verb + a finite verb, NUbvwU †R‡b ivL) and (vi) Complex verb (a noun/adjective/resounding indeclinable + verb, ZvRgnj `k©b Kijvg).

Indeclinable
Indeclinable are of four kinds. They are: (1) Conjunction (Ges), (2) Interjection (evn !), (3) Post position (R‡b¨, Kv‡Q) and (4) Resounding (fb fb).

4. Structures of Bangla Sentences

Structurally, there are three types of Bangla sentences:

a. Simple Sentence (mij evK¨),
b. Complex Sentence (RwUj evK¨),
c. Compound Sentence (‡hŠwMK evK¨).

Each of these has been defined using clause. Clause (LÛevK¨) is the subpart of a Bangla sentence that has a meaning. There are two types of Bangla clause:

a. Principal clause (cÖavb LÛevK¨),
b. Subordinate clause (Avwk^{a}Z LÛevK¨)

Simple Sentence (mij evK¨): A simple sentence is formed by an independent clause or principal clause. Example: kv‡n` ¯‹z‡j hvq . Complex Sentence (RwUj evK¨): A complex sentence is formed by one principal clause and one or more subordinate clause(s). Example: hw` meyR evwZ R¡‡j Zvn‡j iv¯ v cvivcvi K‡iv .
Compound Sentence (‡hŠwMK evK¨): A compound sentence is formed by two or more principal clauses joined by an indeclinable/conjunctive (Ae¨q c`). Example: Aw_wZe„›` mfvq Dcw¯’Z nb Ges mfvcwZ Kvh©µg ïiy K‡ib
Now, a simple sentence can have two parts:
a. Subject (D‡Ïk¨), b. Predicate (we‡aq)
Subject: There are two types of subjects:

a. Simple Subject (mij D‡Ïk¨),
b. Expanded Subject (m¤ú^{a}mvwiZ D‡Ïk¨)

Predicate: There are two types of predicates:

a. Simple Predicate (mij we‡aq),
b.Expanded Predicate (m¤ú^{a}mvwiZ we‡aq)

4.1 Structure of Bangla Complex Sentences

In Bangla complex sentences, there is one principal clause or independent clause and one or more dependent or subordinate clause. These two types of clauses are combined with subordinator and/or corresponding subordinator complement to express complex sentences. For example, a complex sentence "*jadi druto cholo tahole sikgrhoi gontobbe pouchate parbe* (hw` `^{a}yZ P‡jv Zvn‡j kxN^{a}B Mš ‡e¨ †cŠuQ‡Z cvi‡e)". Here, Subordinate clause: *jadi druto cholo* (hw` `^{a}yZ P‡jv). Principal clause: *tahole sikgrhoi gontobbe pouchate parbe* (Zvn‡j kxN^{a}B Mš ‡e¨ †cŠuQ‡Z cvi‡e).
Subordinator: *jadi* (hw`)
Subordinator complement: *tahole* (Zvn‡j)
In complex sentences, above subordinators usually appear before the subordinate clause and subordinator complements are added at various positions in the principal clause [4].

4.2 Bangla Compound Sentences

When two or more independent clauses are connected by preposition (Ae¨q) and thus construct a single sentence, then the sentence is said to be compound sentence [7]. As for example,

Nirbachon hobe ebong gonotontro prothisto hobe (wbev©Pb n‡e Ges MYZš cÖwZwôZ n‡e).
The following prepositions (Ae¨q) are generally used to connect the independent clauses: *o* (I), *ebong* (Ges), *noile* (bB‡j), *fale* (d‡j) etc. The syntactic structure for compound sentence can be established as:
sentence → clause + preposition (Ae¨q) + sentence | clause
Depending on intonation (¯^ifw½), Bangla sentences can be divided into five categories:
a. Assertive Sentence (wee„wZg~jK evK¨),
b. Interrogative Sentence(cÖkœm~PK evK¨),
c. Exclamatory Sentence (we¯§qm~PK evK¨),
d. Optative Sentence (B”Qvm~PK evK¨),
e. Imperative Sentence (Av‡`k evPK evK¨).

4.3 Basic Structural Difference Between Bangla and English

Major differences have been found between Bangla and English languages. Following are the structural differences between Bangla and English:

• The basic sentence pattern in Bangla is subject+object+verb (SOV), whereas in English it is-subject+verb+object (SVO). Example:
R: *cheleti (S) guitar (O) bajai (V)*
B: †Q†jwU MxUvi evuRvq
E: The boy (S) plays (V) guitar (O)

• Bangla is more or less free order structure. Mainly inflection of the components shows relations between the various components of the sentence. Position changes of components normally change the emphasis of an utterance and not the basic meaning.
Example: *badol(S) shuli-ke(O) bhalobash-e(V)* (ev`j wkDwj‡K fvjev‡m) has the same meaning of the sentence *shuli-ke(O) badol(S) bhalobash-e(V)* (wkDwj‡K ev`j fvjev‡m)

• English is a positional language. So it is relatively of fixed order. Relations between various components of the sentence are mainly by the relative positions of the components. Example: "Rahim killed Karim" is quite different from "Karim killed Rahim"

• In English, the modifiers of a subject/object/verb can occur both before and after the object. As an example,

adjectives usually precede nouns, whereas preposition phrases usually follow nouns. Bangla modifiers are followed by subject/object/verb [7]. Example:
R: *Bijoy dibosh-er prathom probhat*
B: weRq w`e‡mi cÖ_g cÖfvZ
E: The first morning of Victory day

5. Bangla Syntax Analysis

5.1 The Proposed Model

Recent propositions to build a Bangla parser do not cover the wide category of ways to form Bangla sentences. For example, the breaking of a sentence into NP and VP may be well suited for simple sentences like, †m wbqwgZ Abykxjb K‡i [Se niomito onushilon kore / He practices regularly]", but sentences that are complex or compound cannot be parsed by this grammar. This is why we have proposed the starting rule to parse all types of sentences, including complex, compound, exclamatory and optative sentences, which were not considered before [1].

Secondly, according to Bangla intonation (¯^ifw½), sentences can be divided into five broad categories. The proposed grammar allows parsing all these types in a new approach.

The parser should accept all Bangla sentences that are syntactically correct, parse each of them using the rules of our proposed grammar. **Fig. 1** illustrates the proposed model.

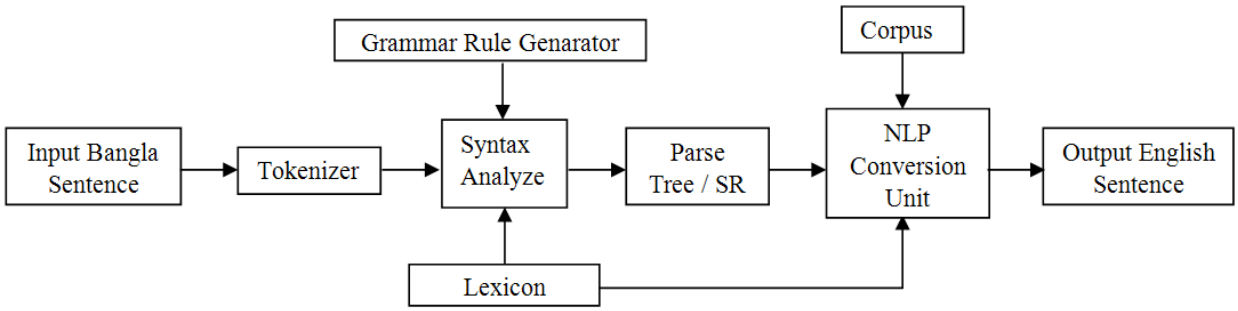

Fig. 1 Block diagram of Bangla parser.

Input Bangla Sentence: The source language sentence is taken as input of the parsing system.
Tokenizer: Tokenizer is the program module that accepts a sentence to be parsed as an unbroken string, breaks into individual words called Tokens. The Token is then checked into the lexicon for validity. The output of Tokenizer is matches with grammatical rules of Bangla. If a rule whose right-hand side matches with a token, then the token is assigned with the appropriate phrases category [1]. The output of the Tokenizer of the input sentence **"Mayeti Gan Gaye Shokolke Mughdho Koreche"** is as follows [1] [2]:

TOKEN=("Maye", "Ti", "Gan", "Gaye", "Shokol", "Ke", "Mughdho", "Kor" "Eche").

Syntax Analyzer/Parser: In this research, the output of Tokenizer is first search in the Lexicon, if the word/token is found then the token/word is matched with grammatical rules of Bangla Grammar Rule Generator. If a rule whose right-hand side matches with a token, then the token is assigned with the appropriate phrases category (parts of speech).
Grammar Rule Generator: The most common way to represent grammar is as a set of production rules which says how the parts of speech can put together to make grammatical, or "well- formed" sentences.
Lexicon: In this paper, the lexicon contains the priori tag and suffix for each word. It has two parts: a *rootform lexicon* and a *suffix lexicon*. During the lookup of a word in the lexicon, the *rootform* lexicon is searched first. If the word is found there, the corresponding phrases category is returned. If it fails, the *suffix lexicon* is searched next and for the rest part of that word, root lexicon is searched. This is done until reach to the beginning of the word from ending point. The advantages of only maintaining root forms in the *lexicon* are: generalizations are captured, knowledge is localized and hence more easily maintained, and new word forms are predicted [8]. So, it is clearly desirable to have only headwords in an MT dictionary. This will save time, space, and effort in constructing entries. **Fig. 2** illustrates a sample lexicon for Bangla parsing.

Noun → শিক্ষা education | জাতি nation | মেরুদন্ড backbone | ভাত rice | মাঠ field | ফুটবল football | রহিম rahim |
Pronoun → তুমি you | সে he | তোমার you |
Inflection → ই is the | এর of the | র of the | কে to | টি the | এ to |
Finite verb → হয় happen | যায় goes | যায় go | খায় eats | কর do | খেলে plays |
Infinite verb → করা to do | বলা to say |
Indeclinable → না no |
Adverb → পরাজিত loser | কোথায় where |
Conjuction → কিন্তু but | এবং and |
Additive word → যেখানে where | সেখানেই there | যাহা what is | তাহা that |
Interogative → কী what |
Adjective →সুন্দর beautiful | মনোরম wonderful | সহজ easy | ভাল good |
Preposition → জন্য for |

Fig. 2 Sample lexicon for Bangla parsing.

Parse Tree/SR: Parsing is a process of finding a parse tree for a given input string. That is, a call to the parsing function PARSE, such as
PARSE ("The rose smells wonderful")
should return a parse tree with root S whose leaves are "The rose smells wonderful" and whose internal nodes are non-terminal symbols. In linear text, we write the tree as …
[S : [NP : [Article : **the**] [Noun : **rose**]] [VP : [Verb : **smells**] [Adjective : **wonderful**]]] **Fig. 3** shows a parse tree for a sample bangle sentence.

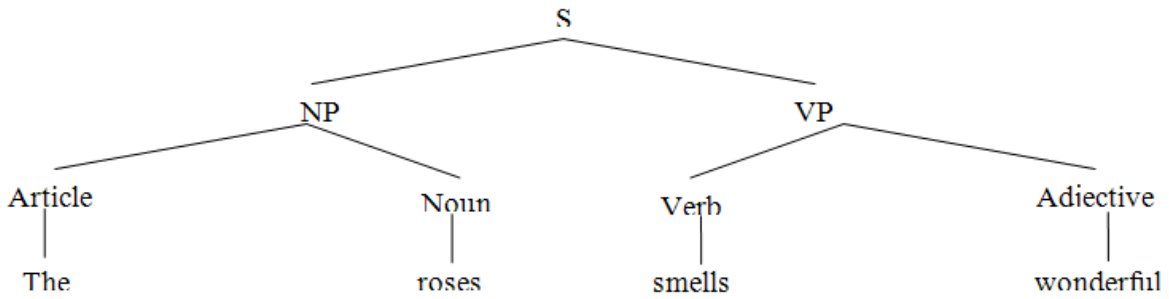

Fig. 3 Parse tree for the sentence "The rose smells wonderful".

NLP Conversion Unit: In this proposed system, an input sentence is analyzed and converted into a parse tree/structural representation (SR). Once the SR is created for a particular sentence it is then converted to corresponding English sentence by NLP conversion unit. For conversion, the NLP conversion unit takes help of the Corpus[8].

Corpus: To implement this Bangla MT, a corpus of large amount of aligned English sentences is used as training corpora. This is the basic training corpus used to train the alignment template Language Model. The language model uses only the English sentences to set the grammatical structure of the expected English sentence as output.

Output English Sentence: The produced English language sentence is the output of the NLP Conversion Unit.

5.2 Basic Rules to Parse a Sentence*

1. Sentence → Simple sentence | Complex sentence | Compound sentence; 2. Simple sentence → Principle clause; 3. Complex sentence → Subordinate part + Additive word + Principal clause | Principal clause + Additive word + Subordinate part; 4. Subordinate part → Subordinate clause | Subordinate clause + Additive word + Subordinate part; 5. Subordinate clause → Additive word + Principal clause; 6. Additive word → Indeclinable | Null; 7. Compound sentence → Principal clause + Additive word + Compound part; 8. Compound part → Principal clause | Compound sentence; 9. Principal clause → Subject + Predicate; 10. Subject → Simple subject | Expanded subject; 11. Predicate → Simple predicate | Expanded predicate; 12. Simple Subject → Actor (KZ©„c`); 13. Actor → Noun + Inflection | Pronoun + Inflection | Implicit (Dn¨) Actor; 14. Pronoun → Person ; 15. Person → FP | SP | TP ; (Example -- FP - aami , aamraa) 16. SP → SPH | SPNH | SPP ; (Example --

SPH- aapni , aapnaara ; SPNH - tumi , tomraa ; SPP - tui , toraa) 17. TP → TPH | TPNH ; (Example -- TPH - tini , taaraa ; TPNH - shey , taaraa) 18. Implicit Actor → Null; 19. Expanded Subject → Sub-expander + Subject; 20. Sub-expander → Adjective | Adjective + Infinite verb | Adjective clause | Relative part (m¤^Ü c`/c`mgwó) | Relative part + Adjective | Adverbial clause; 21. Relative part → Noun + Gi (er) | Pronoun + Gi | Adjective + Gi , 22. Simple predicate → Verb clause | Implicit verb; 23. Implicit verb → Null; 24. Expanded predicate → Pre-expander + Verb clause; 25. Pre-expander → Adverb | Adverb + Adverb | Adverb + Object (Kg©c`) | Adjective + Object | Adjective expander (we‡kl‡Yi we‡klY) + Adjective + Object | Object | Adverbial clause; 26. Object → Noun | Pronoun | Relative part + Noun | Relative part + Pronoun | Null; 27. Verb clause → Infinite verb + Finite verb | Finite verb | Implicit verb | Infinite verb + Finite verb + Indeclinable | Finite verb + Indeclinable(Ae¨q c`); 28. Indeclinable → bv (na) | Other;

* The '→' sign means the phrase "can have the form of", the '|' sign indicates an alternative rule for the left-side term and the '+' sign means join of two terms of a sentence.

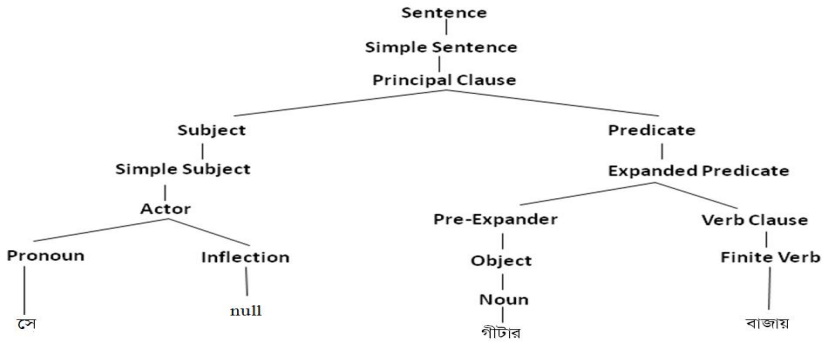

Fig. 4 Parse tree for a Simple Bangla sentence.

6. Implementation of the Proposed Model

6.1 SR Generation using Parsing Methodology

This research performs the SR generation of Bangla sentence by using parsing methodology through the following ways:

1. Take the input Bangla sentence.
2. Calculate total number of Bangla tokens.
3. For each Bangla tokens do the following:
   Find out the equivalent parts of speech and English meaning of the token from Lexicon.
4. Build the corresponding parse tree/SR using grammar rule generator.
5. Output the equivalent English sentence using Bangla to English converter.

6.2 Language Modeling

For implementing the language model, bi-gram model is considered. The flowchart for the bi-gram model is shown in **Fig. 5** and **Fig. 6**, respectively. First, the exact translation of each Bangla word is chosen from the Lexicon and initially an English sentence is constructed. Then the probability of each English word is calculated to come first of the English sentence by training the corpus. The exact word is chosen among these probabilities for which the probability is maximum. After this, from the rest of the words the probability is calculated to come next and so on. And again from these probabilities the next word is chosen for the highest probability. In this way the final English sentence is decorated which follows the training corpus.

6.3 Bangla Translator

After executing the above procedure (total procedure and code not shown) it is possible to translate a Bangla sentence into English. It is perceptible that if the translation of each word is found, the final English sentence would be formed according to the corpus this machine is trained.

7. Experimental Results and Performance

In order to justify the effectiveness of this method, several experiments were conducted. **Fig. 7** illustrates the snapshot of the implemented method. Results for different types of sentences are furnished in Table 1.

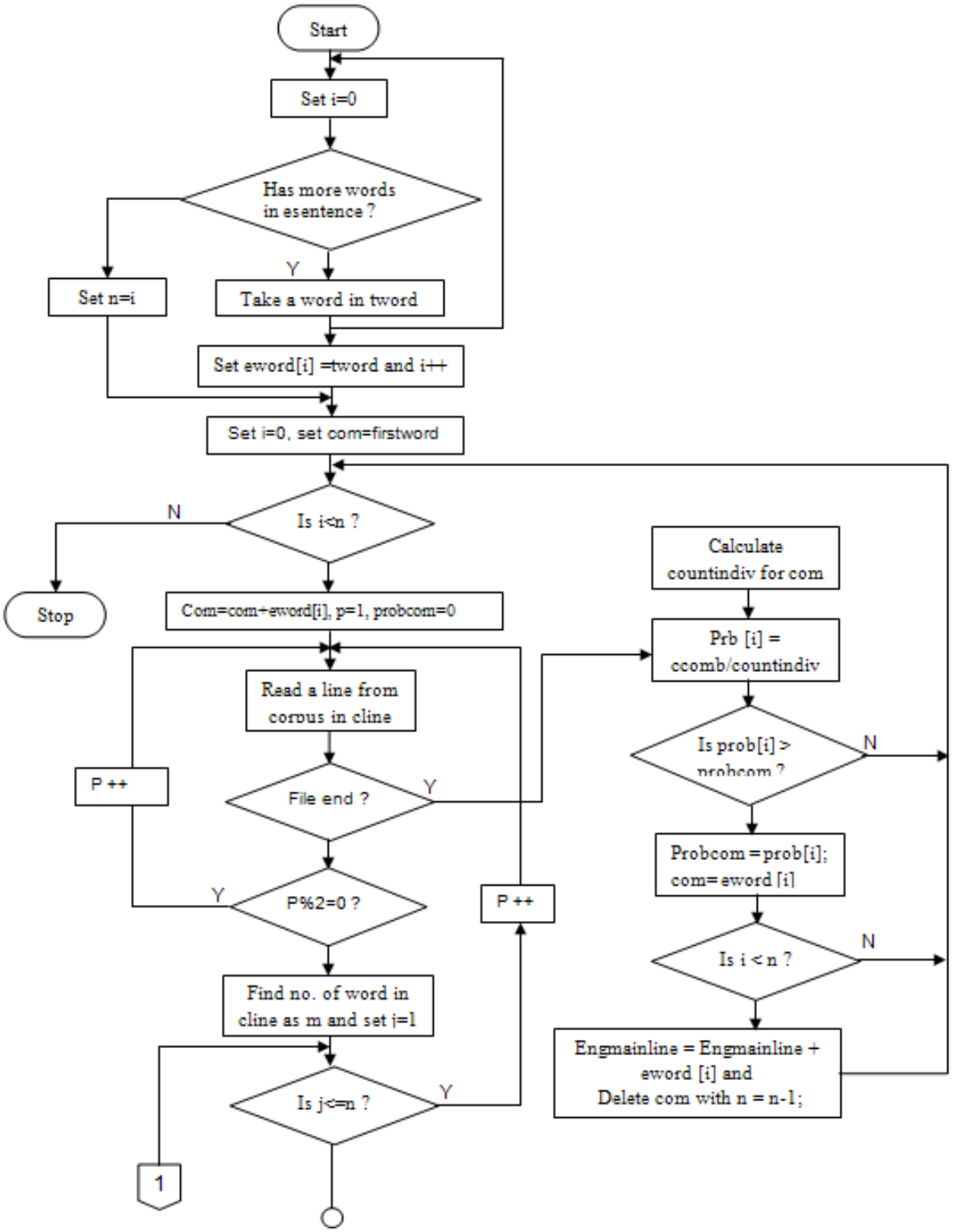

Fig. 5 Flow-chart of Bi-gram model.

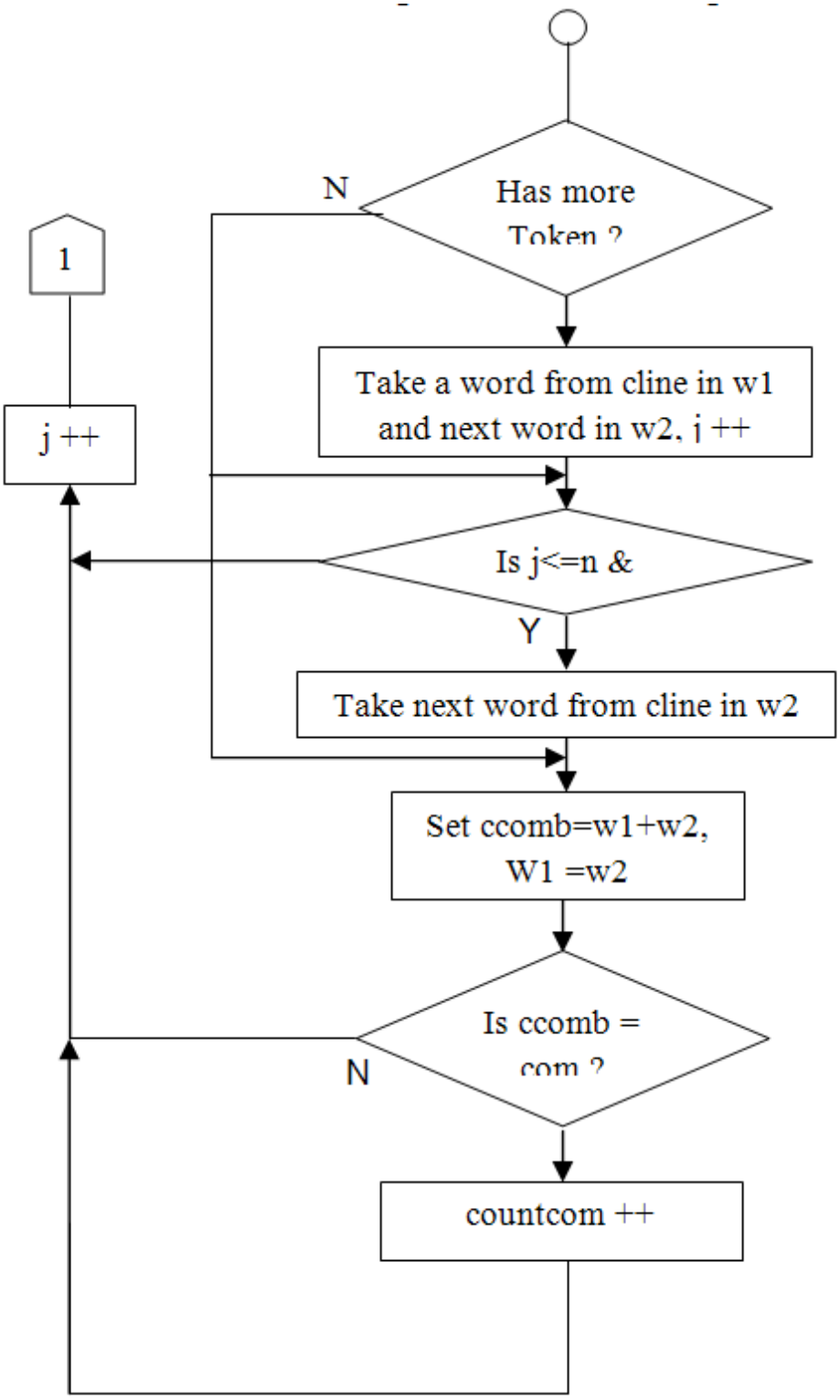

Fig. 6 Flow-chart of Bi-gram model (continue).

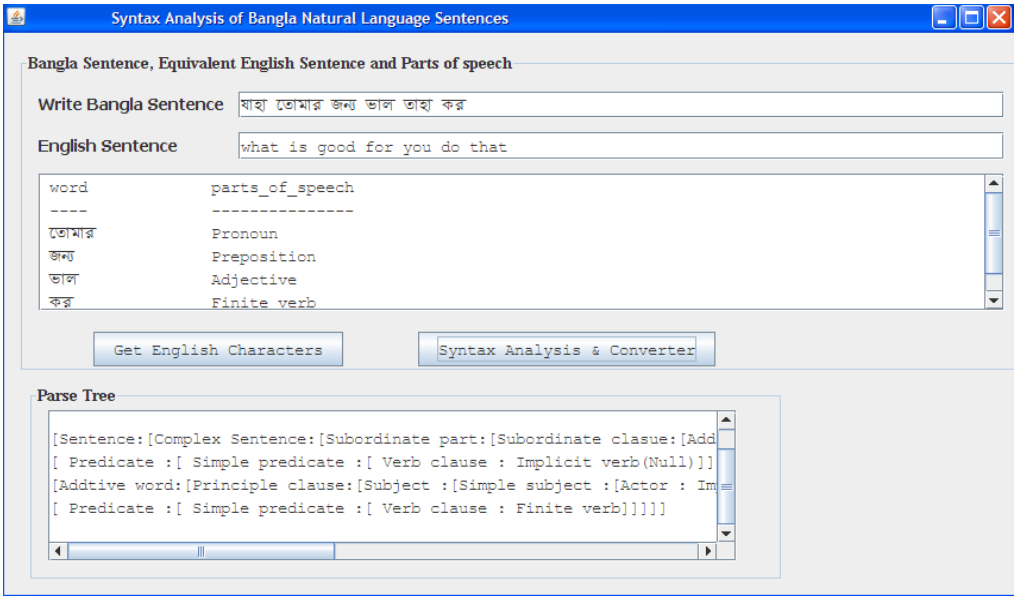

Fig. 7 Sample output of the program for the complex sentence.

Table 1: Success rate for different types of sentences

Type of Sentences	Total no. of sentences	Correctly performed syntax analysis & machine translation	Success rate (%)
Simple	450	420	93.33
Complex	540	500	92.6
Compound	420	385	91.67

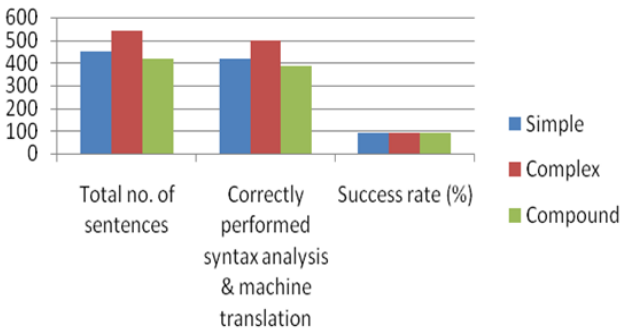

Fig. 8 Success rate for different types of sentences

8. Conclusion

This paper proposes a technique to parse Bangla sentences in a new approach using context-sensitive grammar rules and then convert them to English by NLP conversion unit. The principal goal was to design a parser that is capable of accepting all types of Bangla sentences, from the structural viewpoint. Our proposed parser is based on context-sensitive grammar rule. The major contributions and refinements presented in this paper over the previous works are as follows:

- An innovative parser module for Bangla NL.
- Entry of different words of Bangla language in a lexicon.
- New types of grammars called context-sensitive grammar for Bangla are implemented in this paper.
- How context-sensitive grammar phrase structure rules can be implemented by applying bottom-up approach to parse the simple, complex and compound sentence of Bangla NL has been shown.
- Output of the parser module represented as a list that can be represented and manipulated very easily within a computer.
- Various limitations of context-sensitive grammar to parse the Bangla sentences.

Modifying the context-sensitive grammar rule, we have designed a more powerful parser for Bangla sentences. We have presented a scheme that is effective for transfer and generation of almost all type of sentences for Bangla-English machine translation. The present work extends

and builds upon earlier works. The structural representation that has been built can cover the maximum simple, complex and compound sentences. But the sentences composed of
idioms and phrases are beyond the scope of this paper. Also mixed sentences are of out of the discussion. All the requirements are not being met by these simplified grammars. We are optimistic about future research in this regard.

References

[1] M. M. Hoque and M. M. Ali, "A Parsing Methodology for Bangla Natural Language Sentences", Proceedings of International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, pp. 277-282 (2003).

[2] L. Mehedy, N. Arifin and M. Kaykobad, "Bangla Syntax Analysis: A Comprehensive Approach", Proceedings of International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, pp. 287-293 (2003).

[3] S. Russell and P. Norvig, Artificial Intelligence : A Modern Approach, 2nd Edition, Pearson Education publisher, New York, 2003.

[4] S. K. Chakravarty, K. Hasan, A. Alim, "A Machine Translation (MT) Approach to Translate Bangla Complex Sentences into English" Proceedings of International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, pp. 342-346 (2003).

[5] H. Seddiqui, M. S. Rana, A. A. Mahmud, T. Sayeed, "Parts of Speech Tagging Using Morphological Analysis in Bangla", Proceedings of International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, pp. 274-276 (2003).

[6] T. H. Nipa, M. H. Roshid, S. M. Masum, M. Ali, "Bangla and English Morphological Rules for Machine Translation Dictionary", Proceedings of International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, pp. 391-394 (2003).

[7] K. D. Islam, M. Billah, R. Hasan and M. M. Asaduzzaman, "Syntactic Transfer and Generation of Complex-Compound Sentences for Bangla-English Machine Translation", Proceedings of International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, pp. 321-326 (2003).

[8] M. G. Uddin, M. Murshed, M. A. Hasan, "A parametric approach to Bangla to English Statistical Machine Translation for complex Bangla sentences -Step 1", Proceedings of International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, pp. 529-534 (2005).

Md. Musfique Anwar completed his B.Sc (Engg.) in Computer Science and Engineering from Dept. of CSE, Jahangirnagar University, Bangladesh in 2006. He is now a Lecturer in the Dept. of CSE, Jahangirnagar University, Savar, Dhaka, Bangladesh. His research interests include Artificial Intelligence, Neural Networks, Image Processing, Pattern Recognition, Software Engineering and so on.

Mohammad Zabed Anwar received his B.Sc (Hons) in Electronics and Computer Science and M.S. in Computer Science and Engineering from Dept. of CSE, Jahangirnagar University, Bangladesh. He successfully completed his post-graduate project on Natural Language Processing under the supervision of Dr. Md. Al-Amin Bhuiyan. He is now working as a Database Administrator in a renowned commercial bank in Dhaka, Bangladesh. His main research interests include Natural Language Processing, Artificial Intelligence, Data Warehouse, Data Mining and so on.

Md. Al-Amin Bhuiyan received his B.Sc (Hons) and M.Sc. in Applied Physics and Electronics from University of Dhaka, Dhaka, Bangladesh in 1987 and 1988, respectively. He got the Dr. Eng. degree in Electrical Engineering from Osaka City University, Japan, in 2001. He has completed his Postdoctoral in the Intelligent Systems from National Informatics Institute, Japan. He is now a Professor in the Dept. of CSE, Jahangirnagar University, Savar, Dhaka, Bangladesh. His main research interests include Image Face Recognition, Cognitive Science, Image Processing, Computer Graphics, Pattern Recognition, Neural Networks, Human-machine Interface, Artificial Intelligence, Robotics and so on.