

A New Approach for Identification of Simple Pure Bangla Sentence Type and Checking the Validation in Semantic Manner

Md. Torikur Rahman^{a,*}, F.M.Rahat Hasan Robi^b

^aDepartment of Computer Science & Engineering, Uttara University, Dhaka-1230, Bangladesh

^bDepartment of Computer Science & Engineering, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj-8100, Bangladesh

Abstract

Identifying the sentence type and checking its validity in a semantic manner requires understanding the linguistic structure and meaning of Bangla sentences. It involves analysing the sentence's components, such as subject, verb, object, and other grammatical elements, to determine its type and then evaluating its semantic coherence. To check the validity of a sentence in a semantic manner, we need to consider the coherence and meaning of the sentence within the given context. Semantic validation involves analysing the sentence's components, their relationships, and their overall meaning. Semantic validation can include several checks, such as Grammatical correctness: Verify that the sentence follows the grammatical rules of Bangla and has proper subject-verb agreement, tense consistency, and word order. Meaningful vocabulary: Ensure that the words used in the sentence convey the intended meaning accurately and appropriately. Contextual relevance: Evaluate whether the sentence makes sense within the given context and aligns with the intended message or purpose. Logical consistency: Check if the sentence's components and their relationships are logically coherent and do not contradict each other. Semantic validation in Bangla sentences often requires a profound empathetic of the language and context. It involves analysing the sentence structure, word meanings, and syntactic relationships to ensure the sentence is both grammatically correct and conveys the intended meaning accurately. Now-a-days Bangla sentence type detection and check validation is a growing job. We use semantic feature for sentence type detection and validation check. Semantics is the branch of linguistics that deals with the study of meaning in language. It is concerned with understanding how words, phrases, sentences, and entire texts convey meaning. A word's meaning is not simply a collection of individual letters or sounds; it encompasses the various concepts, associations, and relationships that a word represents. For machine translation correlating words semantic features is useful. This is proficient by endeavouring to break language down into its essential parts. Phrasal categories include the noun phrase, verb phrase, and adjective phrase; lexical categories include noun, verb, adjective, adverb and many others. This paper presents methodology for pure Bangla sentence detection and validation check analysis in semantic manner. Context free grammar and top-down method have been used for the determination of producing interpreted parse tree. Model is tested for a simple sentence of Bangla and test result replicate the effective results for most of the test cases.

Keywords: Bangla Sentence; NLP; Java; Database; Semantic; Lexical

1. Introduction

Natural Language Processing (NLP) is so much difficult rather than other languages. It is difficult because sometimes it doesn't follow the grammar rules strictly. People understand that sentences by using the words meaning, related meaning of sentence, reordering the sequence and get the most closest meaning according to the that situation. The validation of a sentence is also depending on suitability of the sentence. This is measured by the inter-connectivity of the words. Natural languages are huge in size. So, it is so tough to store the relation between all words of a language. To find the type of a sentence and to check the validation of that sentence, here we propose a system. When constructing semantic structures from utterances, it's essential to account for all aspects of meaning within a specific domain. In linguistics and natural language processing, a domain refers to a specific subject area or topic that contextualizes the language being analyzed. Different domains can have their own specialized vocabulary, concepts, and ways of expressing ideas, contact and the task [1]. There is a little work is done on Bangla Language. But no proper work is done. Here we try a new approach. It is not the best but better than other works. We work on detect a

* Corresponding author

E-mail address: torikurrahman@uttarauniversity.edu.bd; torikurrahman@gmail.com

sentence is Imperative or Operative or Exclamatory. In our work, firstly we detect that is valid or not. For this we use sentence structure. We check the relation tables and find the values. Then we check each word's relational value to check validity. Then if the sentence is valid then we go for next phase to detect the sentence type. For detecting the sentence type, we parse the sentence, get its parts of speech, use the semantic rules and finally take the decision.

2. Related Work

Md. Musfique Anwar et al. [2] suggests a method for analysing syntactic structures in Bangla sentences using context-sensitive grammar rules. The approach appears to be designed to accommodate various types of Bangla sentences, including simple, complex, and compound sentences. Additionally, the method includes an NLP conversion unit that interprets the input Bangla sentence and converts it into English. Mohammed Moshikul Hoque and Muhammad Masroor Ali [3] describes a technique for decomposing complex and compound sentences in Bangla. Additionally, it presents a decomposition procedure for Bangla verb phrases, with a focus on extracting necessary information from the auxiliary part. K. M. Azharul Hasan et al. [4] presents a method that is well accepted for parsing a grammar-based language. The method utilizes a top-down parsing approach and incorporates the idea of left factoring to avoid left recursion. Context-Free Grammar (CFG) is employed for modelling the sentence structures of the Bangla language. Mozammel Haque and Mohammad Nurul Huda [5] explores the analysis of Bangla sentences, particularly focusing on the relationship between the subject and verbal inflections. In Bangla, the verbal inflection can change to match the nature of the subject, resulting in multiple syntactically valid verb forms for a single sentence. The authors propose an algorithm for finding semantically valid verb forms for a given subject in Bangla sentences. K. M. Azharul Hasan et al. [6] addresses the significant research problem of semantic error detection in the Bangla language. Due to the rich grammatical, structural, and semantic diversity that Bangla offers, ensuring semantic correctness in sentences becomes essential. The paper focuses on establishing the relationships between subjects and verbs, as well as objects and verbs, in Bangla sentences. Based on this relationship, the authors propose an algorithm for detecting semantic errors in simple Bangla sentences. The algorithm's versatility allows for easy extension to other forms, such as complex and compound sentences. Lenin Mehedy et al. [7] presents a new approach for parsing Bangla sentences using context-free grammar rules. The technique aims to handle various types of Bangla sentences, including complex, compound, exclamatory, and optative sentences. The proposed grammar allows parsing all five categories of sentences based on Bangla intonation. Special emphasis is given to the inflection of verbs, and the paper discusses decomposing verbs and extracting information from their inflection. Sankar et al. [8] shows parsing Bangla language using context free grammar (CFG). They proposed a parser based on Bangla grammar. They used parse table to recognize syntactical mistake of Bangla sentence. Left recursion problem for CFGs is avoided by left factoring. P.P Purohit et al. [9] they proposed a framework for the semantic analyzer that can parse Bangla sentence. To analyze Bangla sentence with semantics they identify several features of Bangla sentence. Chudhuri et al. [10] presents a three-stage approach for named-entity detection in the Bangla language. The approach involves the use of Named-Entity (NE) dictionaries, rules for named-entity, and left-right co-occurrence statistics to identify and extract named entities from text. The experimental results obtained on the Anandabazar Patrika corpus, which is one of the most popular Bangla newspapers, show promising and positive outcomes. Nasreen Akter et al. [11] they try to develop a recognizer for Bangla text. Md Saiful Islam et al. [12] shows a categorization of Bangla sentence using different approach.

3. Proposed System

The rudimentary wedge illustration of the overall process may be showed in Fig. 1. Our system takes Bangla sentences as input and then passes the sentence through a lexical analyzer. The lexical analyzer chop the sentence in smaller parts. The smaller parts of the sentence is tokenized by the lexical analyzer. In this step we find out the word represents which parts of speech from Bangla Grammar. After that the tokens are passed towards syntax analyzer. Syntax analyzer generates the parse tree of our sentence according to given Bangla Grammar. The parse tree is processed by the semantic analyzer, semantic analyzer generates an annotated parse tree. At first, semantic analyzer finds out the word in the dictionary and extract the meaning. A word may have several meaning in several situations. So, the whole sentence should be kept in consideration.

A clear description of the process of semantic processing in natural language processing (NLP) for Bangla sentences. The key steps involved in semantic processing are as follows:

Lexical Analysis: The first step is lexical analysis, where the Bangla sentence is input into the lexical analyzer. The lexical analyzer breaks down the sentence into individual tokens, which are essentially meaningful units or words in the sentence. **Syntax Analysis:** The tokens produced by the lexical analyzer are passed as input to the syntax analyzer. The syntax analyzer uses a predefined grammar to construct a parse tree that represents the syntactic structure of the sentence. **Semantic Analysis:** The parse tree generated by the syntax analyzer becomes the input for the semantic analyzer. The semantic analyzer aims to understand the meaning of the sentence by analyzing the relationships between words and their context. **Word Lookup and Meaning Extraction:** To extract the meaning of individual words, the semantic analyzer looks up each word in a dictionary or lexicon. The dictionary provides the definitions and possible meanings of words. **Multiple Word Meanings:** A challenge in semantic processing arises when some words have multiple meanings or are ambiguous. In such cases, choosing the correct meaning solely based on the word itself may not be sufficient. **Contextual Understanding:** To resolve word ambiguity and identify the correct meaning, the semantic analyzer considers the entire sentence and its context. The meaning of a word can vary depending on its surrounding words and the overall sentence structure. **Semantic Feature of NLP:** The need to analyze sentences as a whole to disambiguate word meanings reflects a key semantic feature of natural language processing. NLP systems must understand the context and relationships between words to achieve accurate semantic understanding.

Semantic processing is a crucial aspect of NLP, as it enables machines to comprehend the meaning of human language. By considering the context and resolving word ambiguities, semantic analysis plays a vital role in various NLP tasks, such as machine translation, sentiment analysis, and question-answering systems. The complexity of semantic processing highlights the challenges and ongoing research in developing more sophisticated and context-aware NLP algorithms.

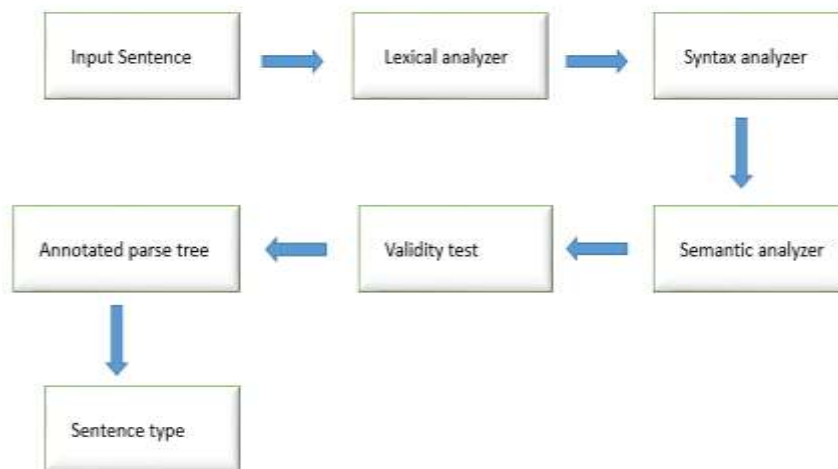


Fig. 1: System Data Flow

1.1. Lexical Semantic Approach

Lexical semantic approach is procedural way of determining meaning of all expressions. In this method the meaning of the whole sentence restatement of the sentence in terms of primitives words and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

1.2. Compositional Semantic Approach

In Compositional Semantic approach procedure meaning obtained from constituent parts. The information structure gained from typical logical expression. The structures shaped by the method are typically prepared as reasonable formula.

1.3. Semantic Features of Bisesso

In Bangla sentences, everything can function as a noun if it represents the name of a person, place, or thing, or if it belongs to a specific class of persons, places, or things. To easily recognize a noun, certain semantic features of the sentence can be analyzed. Nouns in Bangla sentences generally contain the following semantic features: [Animate, Human, Masculine, Singular, Honor, Intelligence]. These features help in identifying and understanding the nature and characteristics of the noun in the sentence. In some sentences, a few semantic features become more dominant than others for the nouns used. For example, in the sentences "Mohila gayok" (female singer) and "purush gayok" (male singer), the significant feature of the two nouns is [Masculine]. "Mohila" is a feminine noun, denoted by Mas [0], while "purush" is a masculine noun, denoted by Mas [1]. The table (Table 1) provides the semantic features (Bisessos) for specific common nouns in Bangla sentences. Each row represents a noun, and each column corresponds to a particular semantic feature.

Table 1: Semantic Feature of Bisesso

| Word | Pranibacok | Monussobacok | Purusbacok | Gayanbcok | Akbocon | Sommanbacok |
|--------|------------|--------------|------------|-----------|---------|-------------|
| Mach | 1 | 0 | X | 0 | 1 | 0 |
| Pakhi | 1 | 0 | X | 0 | 1 | 0 |
| Purus | 1 | 1 | 1 | 1 | 1 | 1 |
| Mohila | 1 | 1 | 0 | 1 | 1 | 1 |

1.4. Semantic Features of Biseson

Pronouns are frequently used in place of nouns in Bangla sentences, especially as subjects. Pronouns play a crucial role in semantic analysis as they convey essential information about the subject or entity being referred to in the sentence. While the semantic features of pronouns are generally similar to those of nouns, there are additional features specific to pronouns that need to be considered for accurate semantic analysis. The feature [Fir] is identified as one of these additional semantic features that impact the analysis of pronouns in Bangla sentences. Two example sentences are provided to illustrate the role of the feature [Fir]. In the sentence "Ami khabo," the verb "khabo" corresponds to the feature [Fir (1)], denoting the first person. The pronoun "Ami" (I) in this sentence also has the feature [Fir (1)], matching the verb. In contrast, in the sentence "tumi khabo," the verb "khabo" corresponds to the feature [Fir (0)], denoting a different person or entity. The pronoun "tumi" (you) in this sentence has the feature [Fir (0)], which is an inverse feature to the verb. A table (Table 2) that lists the semantic features of various pronouns in Bangla, which includes the feature [Fir]. The semantic features of pronouns are crucial for understanding the meaning and context of Bangla sentences. By considering these features, natural language processing systems can accurately interpret the intended meaning and correctly identify the entities being referred to in the sentences.

Table 2: Semantic Feature of Biseson

| Word | Pranibacok | Monussobacok | Gayanbcok | Sommanbacok | Prothombacok |
|--------|------------|--------------|-----------|-------------|--------------|
| Khabo | 1 | 1 | 1 | 0 | 1 |
| Jai | 1 | 1 | 1 | 0 | 1 |
| Jaben | 1 | 1 | 1 | 1 | 0 |
| Ghumao | 1 | 1 | 1 | 0 | 1 |

1.5. Semantic Features of Sorbonam

In Bangla sentences, adjectives are words that represent qualities or characteristics of the subject. When the subject of a sentence is accompanied by an adjective, the adjective becomes the main focus of semantic analysis. Adjectives in Bangla sentences generally have the following common semantic features: [Animate, Human, Masculine]. These

features help in verifying the semantic correctness of sentences. In some cases, adjectives need to have proper and significant semantic features to ensure the sentence's semantic correctness. The features [Ami (1), Hum (1), Mas (x)] are given as examples. For example, the adjectives "sahosi" (clever) and "Valo" (good) have the semantic features [Ami (1), Hum (1), Mas (x)], which indicate that they are suitable for human subjects but not for non-human subjects like "garu." In a sentence, "sundory" can be used as an adjective for a feminine human subject, as it has the semantic feature [Mas (O)]. The table (Table 3) provides the semantic features (Sorbonams) for specific commonly used adjectives in Bangla sentences. Each row represents an adjective, and each column corresponds to a particular semantic feature.

Table 3: Semantic Feature of Sorbonam

| Word | Pranibacok | Monussobacok | Purusbacok |
|--------|------------|--------------|------------|
| valo | 1 | 1 | X |
| sot | 1 | 1 | X |
| kharap | 1 | 1 | X |
| sundor | 1 | 1 | 1 |

1.6. Semantic Features of Kria

In Bangla sentences, verbs are critical as they are semantically related to the subject and play a dominant role in determining the validity and meaning of the sentence. Semantic feature analysis is an approach used to enhance lexical retrieval by improving access to the semantic network, particularly for verbs in this context. Verbs have their own specific semantic features that contribute to the overall meaning of the sentence. These features include [Ani (Animate), H (Human), Mas (Masculine), Adu (Adult), Ho (Honorific), Int (Intelligence)]. The example sentence "Tableti portechi" is analyzed to verify its semantic validity. The verb "portechi" (read) has the semantic features [Ani (1), H (1), Mas (X), Adu (X), Ho (0), Int (X)]. On the other hand, the noun "Table" has the semantic features [Ani (0), H (0), Mas (X), Adu (0), Ho (0), Int (0)]. The mismatch in the feature [Ani] (Animate) indicates that the sentence is semantically wrong, as a non-animate object like a table cannot read itself. The table (Table 4) provides the semantic features (Kria) for specific verbs in Bangla sentences. Each row represents a verb, and each column corresponds to a particular semantic feature.

Table 4: Semantic Feature of Kria

| Word | Pranibacok | Monusobacok | Gayanbcok | Sommanbacok | Prothompurus |
|----------|------------|-------------|-----------|-------------|--------------|
| khai | 1 | X | X | 0 | 1 |
| hattese | 1 | X | 1 | 1 | 0 |
| portese | 1 | 1 | 1 | 1 | 0 |
| ghurtese | 1 | 1 | 1 | 0 | 0 |
| coltese | 1 | X | 1 | 0 | 1 |

1.7. Semantic Features of Kriabiseson

Adverbs in Bangla sentences represent the situation of the verb. There are three main kinds of adverbs: Adverbs of Time, Adverbs of Place, and Adverbs of Manner. These adverbs are crucial for understanding the semantic meaning of Bangla sentences. The two examples provided are "Se agamikal Asbe" (He will come tomorrow) and "Se gotokal asbe" (He will come the day after tomorrow). Both sentences have two adverbs, "agamikal" and "gotokal," which carry specific meanings. Adverbs play a role in semantically validating the sentences. In the first example, the adverb "agamikal" has a semantic feature that is similar to the semantic feature of the verb in the sentence. This similarity makes the first sentence semantically valid. However, in the second example, the adverb "gotokal" has a semantic feature that indicates it should come with the past tense. Since the verb is in the future tense, the second sentence is not semantically valid. The table (Table 5) provides the semantic features (Kriabisesons) for specific adverbs in Bangla sentences. Each row represents an adverb, and each column corresponds to a particular semantic feature.

Table 5. Sematic Feature of Kriabiseson

| Word | Monussobacok | Vobissot | Bortoman |
|----------|--------------|----------|----------|
| agamikal | 1 | 1 | 0 |
| gotokal | 1 | 0 | 0 |
| ajkal | 1 | 0 | 1 |

1.8. Semantic Features of Obboy

In Bangla sentences, certain postfixes are used with nouns and quantifiers in postfix notation. When used with quantifiers, the noun phrase becomes plural, and when used with the right side of the noun, the noun phrase becomes singular. The examples of postfixes provided are "ti," "ta," "jon," "khana," and "khani." When these postfixes are used with quantifiers, the noun phrase becomes plural, and when used with the right side of the noun, the noun phrase becomes singular. The table (Table 6) provides the semantic features (Obboys) for specific postfixes in Bangla sentences. Each row represents a postfix, and each column corresponds to a particular semantic feature. Understanding these semantic features is essential for determining the plurality and singularization of noun phrases in Bangla sentences when postfixes are used.

Table 6: Sematic Feature of Obboy

| Word | Pranibacok | Monusobacok | Manbacok | Sonkhabacok | Bissoybacok |
|-------|------------|-------------|----------|-------------|-------------|
| ti | X | X | 0 | 0 | 0 |
| ta | X | X | 0 | 0 | 0 |
| jon | 0 | 1 | 1 | 1 | 0 |
| khana | 0 | 0 | 0 | 1 | 1 |
| khani | 0 | 0 | 0 | 1 | 1 |

4. Implementation Detailes

This is a complex process to implement semantic rules and other consideration in compact module. We successfully build a prototype module that can efficiently find out the sentence type, its parts of speeches, its meaning is valid or not. In our module we used JAVA for our programming language. All effective data are stored in MYSQL database and some files. GUI of our program is built with JAVA swing. In our processing steps, lexical analysis is done by a function that find out the parts of speech and individual connotation. Then we find out the semantic feature of each detected parts of speech. The whole sentence will go through a function which validate that our input sentence is grammatically valid or not we find out the correlation between all the words and their meaning. If we find perfect association in their meaning, then we predict a decision that our input sentence is grammatically and meaningfully valid or not according to Bengali grammar.

বিশয়সূচক বাক্যঃ বিষয়সূচক শব্দ + কর্তা + কর্ম + ক্রিয়া
আদেশবাচক বাক্যঃ কর্ম + ক্রিয়া
অনুজ্ঞাসূচক বাক্যঃ কর্তা + বিশেষণ / ক্রিয়া বিশেষণ + ক্রিয়া
(নোটঃ কিছু ক্ষেত্রে কর্তা উহ্য থাকে অথবা কর্তার পরে বিধেয় থাকে)
বিবৃতিমূলক বাক্যঃ কর্তা + কর্ম + ক্রিয়া
(বাক্যটি নেতিবাচক হলে ক্রিয়া এর পর না অব্যয় বসে)
প্রশ্নসূচক বাক্যঃ
১। (কেন/কোথায়/কখন/কিভাবে/কি/কাকে/কিসে) ইত্যাদি থাকলে তারপর ক্রিয়া বসে + প্রশ্নবোধক চিহ্ন
২। সর্বনাম + অব্যয় + ক্রিয়া + প্রশ্নবোধক চিহ্ন

Fig. 2: Bangla Structure of Sentence

5. Experiment and Evaluation

Testing is the best way to find out our system successfully works or not. We have collected a number of Bangla sentences. This way we get a large number of test cases. Our system provides successful results for most of the sentences. We provide simple Bangla sentence for our evaluation. All the sentences are placid from Bangla grammar, books, newspapers. For analysing a sentence, we provide a sentence in the Fig. 2 the sentence is "I" which is a valid meaningful Bangla sentence.

input sentence: আমি বই পড়ি

word type: আমি ---> সর্বনাম
বই ---> বিশেষ্য
পড়ি ---> ক্রিয়া

GRAMMAR: সর্বনাম + বিশেষ্য + ক্রিয়া

parse tree: বাক্য (সর্বনাম(আমি) বিশেষ্য(বই) ক্রিয়া(পড়ি))

semantic features: আমি : [প্রানীবাচক(1), মনুষ্যবাচক(1), পুরুষবাচক(X), জ্ঞানবাচক (1), একবচন(1), প্রথম পুরুষ (1)]
বই : [প্রানীবাচক(0), মনুষ্যবাচক(0), পুরুষবাচক(0), জ্ঞানবাচক (X), একবচন(1), সম্মানবাচক(X),]
পড়ি : [প্রানীবাচক(1), মনুষ্যবাচক(1), জ্ঞানবাচক (1), সম্মানবাচক(0), প্রথমপুরুষ(1),]

sentence type: বিবৃতিমূলক বাক্য

validity: সঠিক

Fig. 3: Checking the Validation in Semantic Manner

On the other hand, we provide a sentence, another sentence in Fig. 3 is” 1” which is not a valid sentence with respect to its meaning.

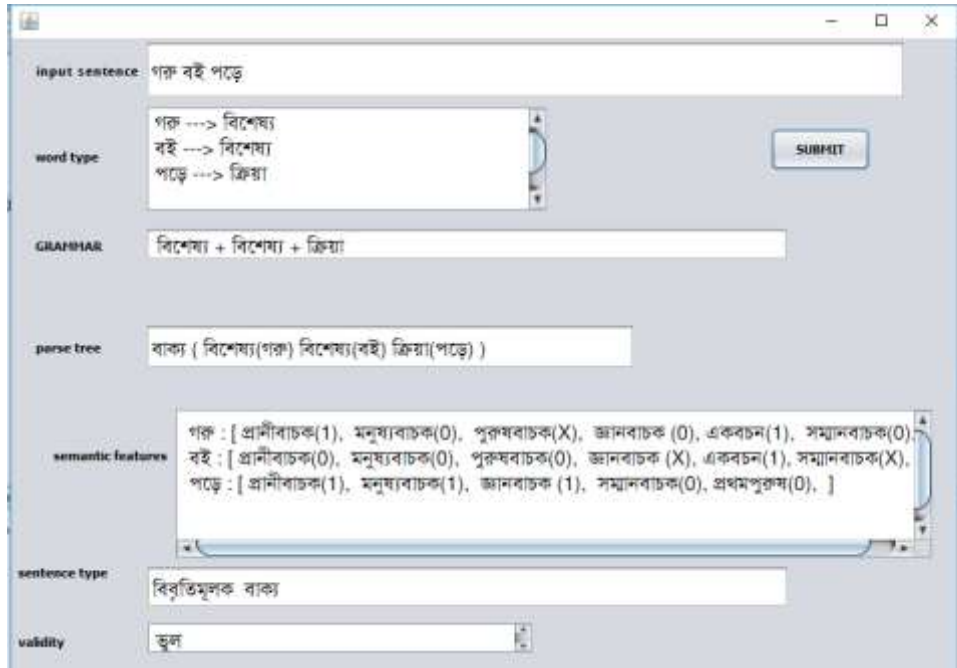


Fig. 4: Sentences checked for validation in Semantic Manner

For the first sentence our system provides that the sentence is semantically valid which provides in the validation field” ” which means the right as well as provide the sentence type” ” which means assertive. It also provides parts of speech of all words in the sentence, grammar rules and annotated parse tree. It also shows the semantic features of each part of speech according to the sentence. Then we validate the sentence that the sentence is grammatically valid or not. If grammatically valid and find out the sentence type. Then we check the semantic validity of the sentence. The first sentence is grammatically and semantically valid. So, our system provides successful analysis of the sentence and provides the correct result. On the other hand, the second sentence in figure. Is grammatically valid but it is not valid semantically. Its meaning is not correct. Our sentence also successfully analyses the sentence. We found that the second sentence is” ” which means assertive. But it is not a meaningful sentence so our output is” ” that means wrong. The best performance is achieved when the sentence is smaller. The accuracy rate decreasing when the sentence word size is in increasing manner.

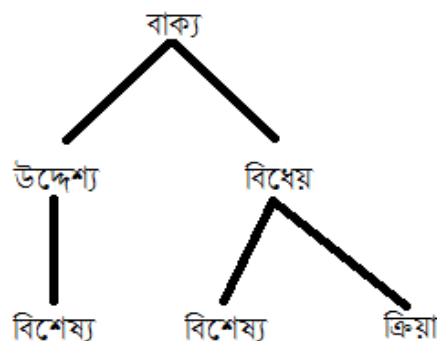


Fig. 5: Annotated Parse Tree.

Table 7: Performance Validation

| No. of input Sentence | Word length | Correctly checked | Error | Accuracy |
|-----------------------|-------------|-------------------|--------|----------|
| 125 | 3 | 125 | 0% | 100% |
| 120 | 4 | 118 | 1.67% | 98.33% |
| 120 | 5 | 109 | 9.167% | 90.83% |

6. Conclusion

Bangla sentence identification is an important issue today. Not only sentence identification but also its type checking and its validation is important task in this era. In this paper we first try to identified the sentence then we try to validate the sentence right or wrong. We take help from pure Bangla grammar to drive our task. We parse the sentence for identification and then check its validation using semantic features. We would like to investigate whether other efficient algorithms would lead to better discovery of type checking and validation. So, processing a Bengali sentence with our algorithm it gives validation, verification and can easily find out which type of Bengali sentence it is. Our system can filter out the invalid sentences. Accuracy of a system express how efficiently the problem domain is handled. the accuracy of the system is pretty best.

References

- [1] Hoque, M.M., Rahman, M.J. and KumarDhar, P., 2007, March. Lexical semantics: A new approach to analyze the bangla sentence with semantic features. In *2007 International Conference on Information and Communication Technology* (pp. 87-91). IEEE.
- [2] Anwar, M.M., Anwar, M.Z. and Bhuiyan, M.A.A., 2009. Syntax analysis and machine translation of Bangla sentences. *International Journal of Computer Science and Network Security*, 9(8), pp.317-326.
- [3] Hoque, M.M. and Ali, M.M., 2004. Context-sensitive phrase structure rule for structural representation of Bangla natural language sentences. *Proceedings of ICCIT, Dhaka*, pp.615-620.
- [4] Hasan, K.A., Mondal, A. and Saha, A., 2010, December. A context free grammar and its predictive parser for bangla grammar recognition. In *2010 13th International Conference on Computer and Information Technology (ICCIT)* (pp. 87-91). IEEE.
- [5] Haque, M. and Huda, M.N., 2016, May. Relation between subject and verb in Bangla Language: A semantic analysis. In *2016 5th International Conference on Informatics, Electronics and Vision (ICIEV)* (pp. 41-44). IEEE.
- [6] Hasan, K.A., Hozaiifa, M. and Dutta, S., 2014, December. Detection of semantic errors from simple Bangla sentences. In *2014 17th International Conference on Computer and Information Technology (ICCIT)* (pp. 296-299). IEEE.
- [7] Mehedy, L., Arifin, N. and Kaykobad, M., 2003. Bangla syntax analysis: A comprehensive approach. In *Proceedings of International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh* (pp. 287-293).
- [8] Sarker, B. and Hasan, K.A., 2013. Parsing bangla grammar using context free grammar. In *Technical Challenges and Design Issues in Bangla Language Processing* (pp. 137-154). IGI Global.
- [9] Purohit, P.P., Hoque, M.M. and Hassan, M.K., 2014, October. An empirical framework for semantic analysis of Bangla sentences. In *2014 9th International Forum on Strategic Technology (IFOST)* (pp. 34-39). IEEE.
- [10] Chaudhuri, B.B. and Bhattacharya, S., 2008. An experiment on automatic detection of named entities in bangla. In *Proceedings of the IJCNLP-08 Workshop on Named Entity Recognition for South and South East Asian Languages*.
- [11] Akter, N., Hossain, S., Sarwar, H. and Rahman, M., 2010. Development of a recognizer for Bangla text: present status and future challenges. In *Character Recognition*. IntechOpen.
- [12] Islam, M.S., Jubayer, F.E.M. and Ahmed, S.I., 2017. A comparative study on different types of approaches to Bengali document categorization. *arXiv preprint arXiv:1701.08694*.