Analysis and Generation of Bengali Case Structure Constructs for Universal Networking Language

Md. Nawab Yousuf Ali	Mohammad Zakir Hossain Sarker	Jugal Krishna Das
Department of CSE	USAID | DELIVER PROJECT	Department of CSE
East West University	House NE(O) 6, Road 92	Jahangirnagar University
Dhaka, Bangldesh	Gulshan-2, Dhaka, Bangladesh	Dhaka, Bangldesh

ABSTRACT
The Universal Networking Language (UNL) deals with the communication across nations of different languages and involves with many different related discipline such as linguistics, epistemology, computer science etc. It helps to overcome the language barrier among people of different nations to solve problems emerging from current globalization trends and geopolitical interdependence. Case structure analysis forms the foundation for any natural language processing task. In this paper we propose the computational analysis of the Bangla case structure and conversion of Bangla sentence to UNL expression. We also provide the implementation of our analysis taking one of the cases as an example.

Keywords
Universal Networking Language (UNL), Bangla grammar, Bangla Word Dictionary, Inflexion, Verbal Inflexion, Bangla Case, Analysis Rule, Enconverter (EnCo)

1. INTRODUCTION
The Universal Networking Language (UNL), which is a formal language for symbolizing the sense of natural language sentences, is a specification for the exchange of information. Currently, the UNL includes 16 languages [1], which are the six official languages of the United Nations (Arabic, Chinese, English, French, Russian and Spanish), in addition to the ten other widely spoken languages (German, Hindi, Italian, Indonesian, Japanese, Latvian, Mongol, Portuguese, Swahili and Thai). In the last few years, machine translation techniques have been applied to web environments. The growing amount of available multilingual information on the Internet and the Internet users has led to a justifiable interest on this area. Hundreds of millions of people of almost all levels of education, attitudes and different jobs all over the world use the Internet for different purposes [2], where English is the main language of the Internet. But English is not understandable for most of the people. Interlingua translation programs are needed to develop. The main goal of the UNL system, which allows users to visualize websites in their native languages, is to provide a common representation for accessing Internet of multilingual websites by the majority of the people over the world. For this common representation, lexical knowledge is a critical issue in natural language processing systems, where the development of large-scale lexica with specific formats capable of being used by distinguished applications, in particular to multilingual systems, has been given special focus. Our goal is to include Bangla in this system with less effort.

So far very little effort has been made to convert Bangla language to UNL expressions. We have been working on this topic from the last 3 years. To convert Bangla sentences into UNL expression we needed to go through the Bangla grammar and UNL very rigorously. Simultaneously we have communicated with the UNL center of the UNDL Foundation. They have made two agreements with us: i) Agreement for entering UNLs and ii) UDS agreement, and provided us user name and password to access their resources and utilities. Then we have started converting Bangla sentences into UNL expression successfully. Although we have already work on Bangla Simple and Compound sentences, but for better understanding of the most of the readers of this paper we have taken a simple affirmative Bangla sentence "আমি ভাত খাই" (Ami Vat Khai) means "I eat rice" and shown how it can be converted into UNL expression.

The organization of this paper is as follow: In Section 2 we describe the Research Methodology/Literature Review, Section 3 has the short description about UNL, Section 4 describes our work. Here we discuss about the dictionary entries and rules, which we have designed and developed to convert for each Bangla cases, section 5 explains how the rules, developed by us, will be applied taking a sentence as an example, section 6 shows the result of our work. In this section we have taken one of the cases as an example although we have worked on all the cases. Finally, Section 7 draws conclusions with some remarks on future works.

2. LITERATURE REVIEW
For converting Bangla sentence to UNL expressions firstly, we have gone through Universal Networking Language (UNL) [3, 4, 5, 6] where we have learnt about UNL expression, Relations, Attributes, Universal Words, UNL Knowledge Base, Knowledge Representation in UNL, Logical Expression in UNL, UNL systems and specifications of Enconverter. All these are key factors for preparing Bangla word dictionary, enconversion and deconversion rules in order to convert a natural language sentence (here Bangla sentence) into UNL expressions. Secondly, we have rigorously gone through the Bangla grammar [7, 8, 9], Verb and roots (Vowel ended and Consonant Ended) [7, 8, 10], Morphological Analysis [10, 11, 12], Primary suffixes [7, 8, 12], Cases and their inflexions [8, 13], construction of Bangla sentence [9] based on semantic structure. Using above references we extort ideas about Bangla grammar for morphological and semantic

analysis in order to prepare Bangla word dictionary (for root, root word suffix etc), morphological rules and enconversion rules in the format of UNL provided by the UNL center of the UNDL Foundation.

3. UNL SYSTEM – IN A NUTSHELL

Although, there is an immense proliferation of information through Internet, it is not accessible to vast multitude of people across nations as most of the resources are in English. To overcome this problem, United Nations launched Universal Networking Language project [10] in 1996. The result of the project is universal networking language (UNL), a language neutral specification, and a universal parser specification [11]. The goal is to eliminate the massive task of translation between two languages and reduce language to language translation to a one time conversion to UNL.

The UNL [12] has been introduced as a digital meta language for describing, summarizing, refining, storing and disseminating information in a machine independent and human language neutral form. This meta-language focuses to express meanings in standardized way. We think that a comprehensive description of UNL specification is necessary though it is available in UNL website. The meaning of native language sentence is expressed in UNL system as a hypergraph composed of nodes connected by semantic relations. Nodes or Universal Words (UWs) are words loaned from English and disambiguated by their positioning in a knowledge base (KB) [10] of conceptual hierarchies. Function words, such as determiners and auxiliaries are represented as attributes to UWs or nodes to provide additional information. The core structure of UNL is based on the following elements [4]:

1. Universal Words
2. Attribute Labels
3. **Relation Labels**
4. UNL Expression
5. **Hypergraph**
6. **Knowledge Base**

4. PROPOSED UNL STRUCTURE OF BANGLA CASE

As in the traditional understanding case (in Bangla Kaarok) in Bangla denotes the relationship of the nominal with the main verb of the clause. The cases are broadly classified into 6 categories [7], each having a finer categorization into sub-types. The traditional case with their inflexions is shown in table 1.

Table I. Bangla Cases and their inflexions

Classical Case vs. Bangla Kaarok	Inflexions (Case maker)	
	Singular	Plural
Nominative case(*Kartri Kaarok*)	0 (null), এ (e), য় (oy), তে (te)	রা (ra), এরা (era)
Accusative case (*Karma Kaarok*)	0 (null), কে (ke), রে (re), এরে (ere)	দিগকে (digoke)
Instrumental case (*Karan Kaarok*)	দ্বারা (dara), দিয়া (dia), কর্তৃক	দের দ্বারা (der dara), দিগের

	(kortik)	দ্বারা (diger dara)
Dative case (*Sampradaan Kaarok*)	0(null), কে(ke), রে(re), এরে(ere), এ(e), য়(oy), তে(et)	দিগকে (digoke)
Ablative case (*Apaadaan Kaarok*)	হইতে (hoite), থেকে (theke)	দের হইতে (der hoite), দিগের হইতে (diger hoite)
Genitive case (*Sambandha pad*)	র (r), এর (er)	দের (der), দিগের (diger)
Case of time-place (*Adhikaran Kaarak*)	এ (e), য় (oy), তে(te)	সকলে (sokole), দের মধ্যে (der modde)

In this paper, an exhaustive realization of the case system with a view to analyze Bangla into UNL has been carried out.

4.1 Nominative case (Kartri Kaarok)

This case is the form of a noun or pronoun used in the subject or predicate nominative. It denotes the agent of the action stated by the verb. For example, "করিম পড়িতেছে", pronounce as *Karim Poritechhe* means "*Karim is reading*". Here subject *Karim* initiates an action. So, agent (*agt*) relation is made between subject "Karim" and verb "read".

The following dictionary entries are needed for conversion the sentence [14]

[করিম] {} "karim (icl>name, iof>person,com>male)(N)

[পড়] {} "read (icl>see>do,agt>person,obj>information) (ROOT,CEND,^ALT)

[ইতেছে] {} ""'(VI, CEND, CEG1,PRS,PRG, 3P)

where, N denotes noun, ROOT for verb root, CEND for Consonant Ended Root, ^ALT for not alternative, VI is attribute for verbal inflexion, CEG1 for Consonant Ended Group 1, PRS for present tense, PRG for progress means continuous tense and 3P for third person.

To convert this sentence into UNL expression morphological analysis is made between "পড়" (por) and "ইতেছে" (itechhe) and semantic analysis is made between "করিম" (karim) and "পড়িতেছে" (poritechhe)

- Rule of morphological analysis:

After applying some right shift rules when verb root "পড়"(por) comes in the LAW and verbal inflexion "ইতেছে"(itechhe) comes in the RAW then following rule is applied to perform morphological analysis discussed in 3.3.1.1

+{ROOT,CEND,^ALT,^VERB:+VERB,-ROOT,+@::}{VI,CEND:::}P10;

- Rule of semantic analysis:

After morphological analysis when noun "করিম" (karim) appears in the LAW and verb "পড়িতেছে" (poritechhe) appears in the RAW the following agent (agt) relation is made between "করিম" (karim) and "পড়িতেছে" (poritechhe) to complete the semantic analysis.

>{N,SUBJ::agt:}{VERB,#AGT:+&&@present,+&&@progress::}P1;

where, NOM denotes the attributes for nouns or pronouns

UNL Expression:

```
{unl}
agt(read(icl>see>do,agt>person,obj>information).@entry.@prese
nt.@progress,karim(icl>name,iof>person,com>male))
{/unl}
```

4.2 Accusative case (Karma Kaarok)

Accusative case is a noun or pronoun or thing on which the subject executes the verb. For example: "আমি   ভাত খাই।" pronounce as "Aami vat khai" means "I eat rice."

The following dictionary entries are needed for the sentence.
[আমি]{} "i(icl>person)"(PRON,HPRON,1P,SG,SUBJ)<B,1, 1>
[ভাত]{} "rice(icl>food)"(N)<B,0, 0>
[খা]{}"eat(icl>consume>do,agt>living_thing,obj>concrete_thing)" (ROOT,VEND,#AGT,#OBJ,VEG1)<B,0,2>
[ই]{}"" (VI,VEND,1P,PRS)<B,0,0>

where, PRON denotes pronoun, HPRON for human pronoun, 1P for first person, SG for singular number and SUBJ for subject of the sentence.

To convert this sentence into UNL expressions a morphological analysis and two semantic analyses to be held [15].

- Rule for morphological analysis:

Morphological analysis between 'খা' (kha) and 'ই' (e) to be held using the following rule detailed in section 4.1
+{ROOT,VEND,^ALT,^VERB:+VERB,-ROOT,+@::}{VI, VEND:::}P10;

- Rules for semantic analyses:

For the above sentence two semantic relations are made. One semantic relation is agent (agt) relation to be held between "আমি" (aami) and "খাই" (khai) using the following rule detailed in section 4.1
>{PRON,HPRON,SUBJ::agt:}{VERB,#AGT:::}P10;

And another relation is between "ভাত" (rice) and "খাই" (eat)
Here, rice is an *object* and relation between rice and eat is the accusative case.
UNL rule for accusative case to perform semantic analysis is (detailed in 3.3.2.2):
>{N::obj:}{VERB,#OBJ:+&@present.&@entry::}P10;

UNL expression:

```
{unl}
agt(eat(icl>consume>do,agt>living_thing,obj>concrete_thing).@
entry.@present,i(icl>person))
obj(eat(icl>consume>do,agt>living_thing,obj>concrete_thing).@
entry.@present,rice(icl>grain>thing))
{/unl}
```

4.3 Instrumental case (Karan Kaarok)
Instrumental case is the thing, tool or method by which the subject of the sentence executes the specified action.

For example:
"সে কলম দিয়ে একটি নোট লিখেছিল।" pronounce as *Se kolom die ekti   note likhechhilo, m*eans "He wrote a note with a pen."
The following dictionary entries are needed for converting the sentence into UNL expression.

[সে] {} "he(icl>person)"(PRON,HPRON,1P, SUBJ, MALE)
[কলম] {} "pen(icl>writing_implement>thing"(N)
[দিয়ে] {} "" (INF, INF2ND, PRON)
[নোট] {} "note(icl>personal_letter)>thing)"(N)
[লিখ্] {}"write(icl>do, agt>person, obj>abstract_thing plc>thing, ins>functional_thing)" (ROOT, CEND)
[এছিল]{} "" (VI, CER, 3P, PST)

where, MALE denotes male, INF is the attribute for inflexion, INF2ND for second types of inflexion and PST for past tense.

In this sentence two morphological and three semantic analyses are taken place:
- Rules for Morphological Analyses:
The following rules are applied to complete morphological analyses between "কলম" (kolom) & "দিয়ে" (die) and between "লিখ্" (likh) & "এছিল" (echhilo) detailed in section 3.3.1.1
 +{N:@::} {INF, INS:::}

+{ROOT,CEND,^ALT,^VERB:+VERB,-ROOT,+@::}{VI, CEND : &@past::}

- Rules for semantic analyses:
First semantic relation is object (obj) relation which is made between "নোট" (note) "লিখেছিল" (likhechhilo) using the following rule discussed in section 3.3.2.2

>{N::obj:}{VERB,#OBJ:+&@past::}P10;

Second semantic relation is made between "কলম" (pen) which is instrument and "লিখেছিল" (wrote) is instrumental case.

Rule for instrumental case to perform semantic analysis is: (detailed in section 3.3.2.2)
>{PRON, HPRON, INS::ins:}{VERB:+&@past::}

Third semantic relation is object (obj) relation to be held between "সে" (he) and "লিখেছিল" (wrote) using the following rule detailed in section 4.1

>{PRON,HPRON,SUBJ::agt:)}{VERB:+&@past.&@entry::}

UNL expression:

```
{unl}
agt(write(icl>do,agt>person,obj>abstract_thing,plc>thing,ins>fun
ctional_thing).@entry.@past,he(icl>person))
obj(write(icl>do,agt>person,obj>abstract_thing,plc>thing,ins>fun
ctional_thing).@entry.@past,note(icl>personal_letter>thing).@in
def)
ins(write(icl>do,agt>person,obj>abstract_thing,plc>thing,ins>fun
ctional_thing).@entry.@past,pen(icl>writing_implement>thing).
@indef)
{/unl}
```

4.4 Dative case (Sampradaan Kaarok)

In Dative case, subject of a sentence does or gives something for or to someone that means a noun or pronoun is in the Dative case when it is used as an indirect object.

For example: "রাজা তাহার পুত্রকে মুকুট দিবে।" pronounce as *Raja tahar putroke mukut dibe* means "The king will give crown to his son"

The following dictionary entries are needed for converting the sentence into UNL expression.

[রাজা]{} "king (icl>sovereign>thing,ant>queen)" (N,3P)

[তাহা]{} "he(icl>person)"(PRON,HPRON,3P)

[র]{} "" (INF, INF6TH,OBJ, POS)

[পুত্র]{} "son(icl>male_offspring>thing,ant>daughter)" (N)

[কে]{} "" (INF, INF2ND, NOM, OBJ, BEN)

[মুকুট]{} "crown(icl>jewelled_headdress>thing)"(N)

[দি]{}"give(icl>do,equ>hand_over,agt>thing,obj>thing,ben>person)" (ROOT,VEND,VEG1,#AGT,#OBJ)

[বে]{} "" (VI,VER,3P,FUT)

where, FUT denotes future tense, NOM indicates noun or pronoun, and POS for possessive.

In the above sentence, three morphological and four semantic analyses are made to convert the sentence into UNL expressions.

- Rules for morphological analyses:

First, second and third morphological analyses to be held between "দি" (di) & "বে" (be), "পুত্র" (putro) & "কে" (ke) and "তাহা" (taha) & "র" (r) to complete the meaning of the words "দিবে" (dibe), "পুত্রকে" (putroke) and "তাহার" (tahar) respectively using the following morphological rules discussed detailed in section 3.3

Rule1: +{ROOT,VEND,^ALT,^VERB:+VERB,-ROOT,+@::} {VI,VEND:@future::}P10;
Rule2:+{N:@::}{INF, 2NDINF, NOM:::}P10;
Rule3: +{PRON,HPRON:@::}{INF,6THINF,NOM:::}P10;

- Rules for semantic analyses:

First semantic relation is object (obj) relation which is made between "মুকুট" (crown) "দিবে" (will give) using the following rule discussed in section 3.3.2.2
>{N::obj:}{VERB,#OBJ:+&&@future::}P10;

Second semantic relation is made between "পুত্রকে" (to son), which is beneficiary and "দিবে" (will give) is dative case.

Rule for dative case to perform semantic analysis is: (detailed in section 3.3.2.2)

>{N,BEN::ben:}{VERB:+&&@future::}

Third semantic relation is possessive (pos) relation to be held between "তাহার" (his) and "পুত্রকে" (to son) using the following rule detailed in section 4.1

>{PRON,HPRON,OBJ::pos:)}{N:::}

Forth semantic relation is agent(agt) relation to be held between "রাজা" (king) and "দিবে" (give) using the following formula detailed in section 3.3.2.2

>{N,SUBJ::agt:)}{VERB:+&&@future,&&@entry::}

UNL expression:

```
{unl}
agt(give(icl>do,equ>hand_over,agt>thing,obj>thing,rec>person).
@entry.@future,king(icl>sovereign>thing,ant>queen).@def)
obj(give(icl>do,equ>hand_over,agt>thing,obj>thing,rec>person).
@entry.@future,crown(icl>jewelled_headdress>thing))
pos(son(icl>male_offspring>thing,ant>daughter),he(icl>person))
ben(give(icl>do,equ>hand_over,agt>thing,obj>thing,rec>person).
@entry.@future,son(icl>male_offspring>thing,ant>daughter))
{/unl}
```

4.5 Ablative case (Apaadaan Kaarok)

It describes the concept of sources of creating, location, position etc. All types of relations having the concept of source are eligible to come into this category.

For example: "পাখিটি খাচা থেকে উড়ে যায়।" pronounce as *Pakhiti khacha theke ure jae* means "The bird flys from the nest"

The following dictionary entries are needed for converting the sentence into UNL expression.

[পাখি]{}"bird(icl>vertebrate>thing)"(N)

[টি]{} ""(ART)

[খাচা]{} "nest(icl>retreat>thing)"(N,CONCRETE)

[থেকে]{} "from(icl>how,obj>thing,plc>uw)"(FRM)

[উড়]{} "fly(icl>do)"(ROOT, CEND,CEG,#FRM,#TO)

[এ]{} ""(INF, 7THINF,CEND)

[যা]{} "go(icl>move>do)"(ROOT,VEND, VEG1,#AGT, #PLC)

[য়]{} ""(VI,VEND,3P,PRS)

In the above sentence, four morphological and three semantic analyses are made to convert the sentence into UNL expressions.

- Rules for morphological analyses:

First, second and third morphological analyses to be held between "যা" (ja) & "য়" (ye), "উড়" (ur) & "এ"(e), "খাচা" (khacha) & "থেকে" (theke) and "পাখি" (pakhi) & "টি" (ti) to complete the meaning of the words "যায়" (jae), "উড়ে" (ure), "খাচা থেকে" (khacha theke) and "পাখিটি" (pakhiti) respectively using the following morphological rules discussed detailed in section 3.3

Rule1: +{ROOT,VEND,^ALT,^VERB:+VERB,-ROOT,+@::} {VI,VEND:@present::}P10;
Rule2: +{ROOT,CEND,^ALT,^VERB:+VERB,-ROOT,+@::}{VI,CEND:::}P10;
Rule 3 : +{N:@::} {INF, FRM :::}P10;
Rule 4: +{N:@::}{INF,NOM:::}P10;

- Rules for semantic analyses:

First semantic relation is object (obj) relation which is made between "খাচা" (nest) and থেকে" (from) using the following rule discussed in section 3.3.2.2
>{N::plc:}{FROM,#PLC:::}P10;

Second semantic relation is from (frm) relation which is made between "খাচা থেকে" (from nest) and "উড়ে যায়" (fly) using the following rule discussed in section 3.3.2.2
> {N, FRM::frm :}{VERB,#FRM:+&&@present ::} P10;

Third semantic relation is agent (agt) relation, made between "পাখিটি" (the bird), and "উড়ে যায়" (fly) using the following formula detailed in section 3.3.2.2

>{NOM::agt:}{VERB:+&&@present.&&@entry::}

UNL expressions:

{unl}
agt(fly(icl>move>do,equ>move_around,plt>place,plf>place,agt>thing),bird(icl>vertebrate>thing).@def)
frm(fly(icl>move>do,equ>move_around,plt>place,plf>place,agt>thing),from(icl>how,obj>thing,plc<uw))
obj(from(icl>how,obj>thing,plc<uw),nest(icl>shell>thing).@def)
{/unl}

4.6 Case of time-place (Adhikaran Kaarak)

This case describes the place, time and topic of the action performed by the sentence. For example, "আমরা ঢাকায় থাকি" pronounce as *Aami dhakae thaki* means "I live in Dhaka."

The following dictionary entries are needed for the sentence.
[আমরা]{}"we(icl>group)"(PRON,HPRON,1P,SG,SUBJ) <B,1,1>
[ঢাকা]{}"dhaka(iof>city)"(N)<B,0, 0>
[য়]{}""(INF,N)<B,0,2>
[থাক]{}"(live(icl>inhabit>be,aoj>living_thing,plc>place)"(ROOT, CEND,#AGT,VEG1)<B,0,2>
[ই]{}""(VI,CEND,1P,PRS)<B,0,0>

To convert this sentence into UNL expressions two morphological analyses and two semantic analyses to be held.

- Rule for morphological analysis:

Morphological analyses between 'থাক' (thak) & 'ই' (i) and "ঢাকা" & "য" to be held using the following rules detailed in section 4.1
Rule1:+{ROOT,CEND,^ALT,^VERB:+VERB,-ROOT,+@::}{VI,CEND,1P:::}P10
Rule2: +{N:@::}{INF,N:::}P10

- Rules for semantic analyses:

First semantic relation is thing with attribute (aoj) relations to be held between "আমরা" (we) and "থাকি" (live) using the following rule detailed in section 4.1
>{PRON,HPRON,SUBJ::agt:}{VERB,#AGT:&&@pl::}P10;

And another semantic relation is place (plc) relation between "ঢাকায়" (in Dhaka) and "থাকি" (live).
Here, Dhaka is a *place* and relation between *place* and *live* is the case of time-place.
UNL rule for *case of time-place* to perform semantic analysis is (detailed in 3.3.2.2):
> {N, PLACE,:plc :}{ V:+&&@present.&&@entry::}

Unl expressions :

{unl}
aoj(live(icl>inhabit>be,aoj>living_thing,plc>place).@entry.@present,we(icl>group).@pl)
plc(live(icl>inhabit>be,aoj>living_thing,plc>place).@entry.@present,dhaka)
{/unl}

5. CONVERSION OF BANGLA SENTENCE INTO UNL

Although we have analyzed all of the above cases in detail for simplicity and better understanding of the readers we have considered a sentence "আমি ভাত খাই" (pronounce as *aami vaat khai*.) meaning, "I eat rice." for conversion process. Assuming that all the words and morphemes of the given sentence are in the dictionary as follows:

[আমি]{}"i(icl>person)"(PRON,HPRON,1P,SG,SUBJ)<B,1,1>
[ভাত]{}"rice(icl>food)"(N)<B,0, 0>
[খা]{}"eat(icl>consume>do,agt>living_thing,obj>concrete_thing)"(ROOT,VEND,#AGT,#OBJ,VEG1)<B,0,2>
[ই]{}"KBIV"(KBIV,VEND,1P,PRS)<B,0,0>

where, #AGT for agent (means agent relation can be made with root 'খা' (kha) , #OBJ for object (like as agent relation), VEG1 means the root is fall in the vowel ended group 1.

EnCo can input either a string or a list of words for a sentence of a native language. A list of morphemes or words of a sentence must be enclosed by [<<] and [>>] [4, 5]. When the sentence is taken into EnCo, it places the sentence head (<<) in the LAW, sentence texts or morphemes or words in the RAW and the sentence tail (>>) in the RCW shown in figure 1.

Figure 1: Initial state of the Analysis Windows and node list

After insertion of the input file (InputFile.txt) with our given sentence the following rules will be applied step by step to complete the conversion process of the sentence to UNL expressions.

Rule 1:
R{SHEAD:::}{PRON,SUBJ:::}P10;
Rule 2:
DR{SUBJ,^blk:blk::}{BLK:::}P10;
Rule 3:
R{PRON,SUBJ:::}{N:::}P10;
Rule 4:
DR{N,^blk:blk::}{BLK:::}P10;
Rule 5:
R{N:::}{ROOT,^VERB:::}P10;
Rule 6:
+{ROOT,VEND,^ALT,^VERB:+VERB,-ROOT,+@::}{VI,VEND:::}P10;
Rule 7:
:{:::}{VERB,VI:-VI,-VEND,-CEND::}P10;
Rule 8:
>{N::obj:}{VERB,#OBJ:::}P10;
Rule 9:
>{HPRON,SUBJ::agt:}{VERB,#AGT:::}P10;

Rule10:
R{SHEAD:::}{VERB,^&@entry:+&@entry::}P10;
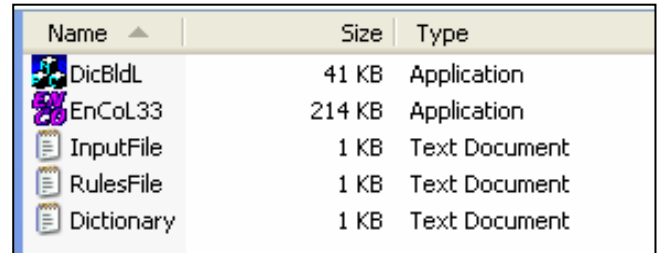Rule 11:
R{VERB:::}{STAIL:::}P10;
Rule 1 is *Right Shift rule* that describes that when sentence head is in the Left Analysis Window (LAW) and word 'আমি' (aami) is in the Right Analysis Window (RAW) then AWs will be shifted to right after rule application. In this situation, the Enco will retrieve the word, 'আমি' (aami) from the Word Dictionary file and remains in the LAW and 'ভাত খাই' (vat khai) will be in the RAW. Rule 2 is *Right Node Deletion rule*, it deletes the right node which is blank space between 'আমি' (aami) and 'ভাত' (vat) and only the noun 'ভাত' (vat) will be placed in the RAW while the verb 'খাই' (khai) will be placed in the Right Condition Window (RCW). Again *right shift rule* (rule 3) is applied to shift the windows to right and *Right Node Deletion rule* (rule 4) is applied to delete the space between 'ভাত' (vaat) and 'খাই' (kaai) so that the word 'ভাত' (vaat) is retrieved from the Word Dictionary and remains in the LAW and the verb 'খাই' (kaai) is divided into root 'খা' (kha) which remains in the RAW and verbal inflexion 'ই' (i) remains in the RCW. Now, again *right shift rule* (rule 5) is applied to place 'খা' (kha) in the LAW and 'ই' (i) in the RAW to perform the morphological analysis. In this situation, Enco retrieves the dictionary entries of 'খা' (kha) and 'ই' (i) from the word dictionary (input file) and will apply the analysis rule (rule 6) that is if there is a vowel ended root (in our example, 'খা') is in the LAW and verbal inflexion (in our example, 'ই') is in the RAW then after applying the rule the two headwords of the left and right nodes are combined into a composite node to complete the morphological analysis of the verb 'খাই' (khai). Rule 7 is *attribute changing rule* used to rewrites the attributes by deleting attributes KBIV, VEND, and CEND the for verb 'খাই' (khai) that remains in the RAW. Now Enco starts semantic analysis between the words of our sentence by applying semantic rules 8 and 9. Rule 8 is *Right Modification Rule* (>), indicates that if noun 'ভাত' is in the LAW and verb 'খাই' is in the RAW then an object relation (obj) is made between them where 'খাই' remains in the RAW and 'ভাত' is deleted. Now the word 'আমি' (aami) comes to the LAW and an agent relation (agt) is made between 'আমি' (aami) and 'খাই' (khai) by applying rule 9, so that 'আমি' (aami) is deleted from the node-list and the verb 'খাই' (khai) remains in the RAW which is the main predicate of the sentence. After that right shift rule 10 is applied to shift the windows to right and *&&entry* attribute is added to the verb as verb 'খাই' (khai) is the main word of the sentence. Finally, rule 11 is applied to place the sentence tail (STAIL) on the LAW to complete the conversion process [4]

6. RESULT

To convert the Bangla sentence "আমি ভাত খাই" (Ami Vat Khai) we have used the following files.
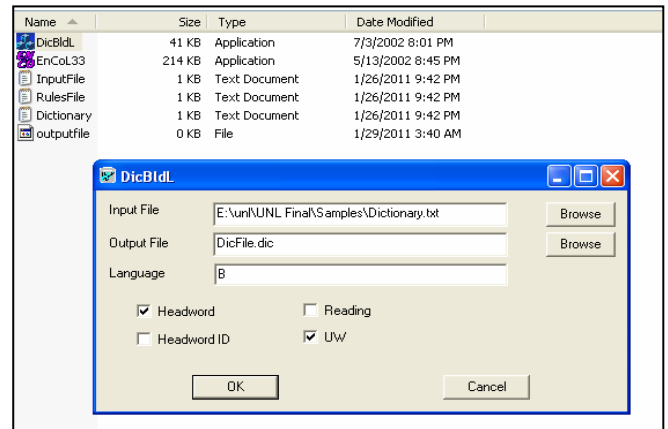1. Input file
2. Rules File
3. Dictionary

We have used an Encoder (EnCoL33.exe) and a dictionary builder file (DicBldL.exe) provided by the UNDL Foundation of UNL center, which we have downloaded from [16].
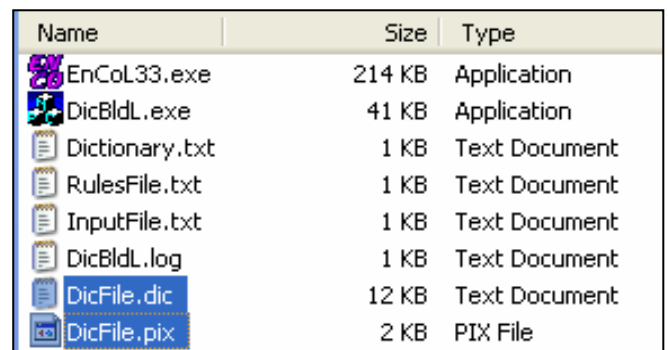Screen 1 shows the files to be used for EnCo

Screen 1: The files used for conversion

Screen 2 shows the way of creating dictionary file using *DicBldL.exe* file

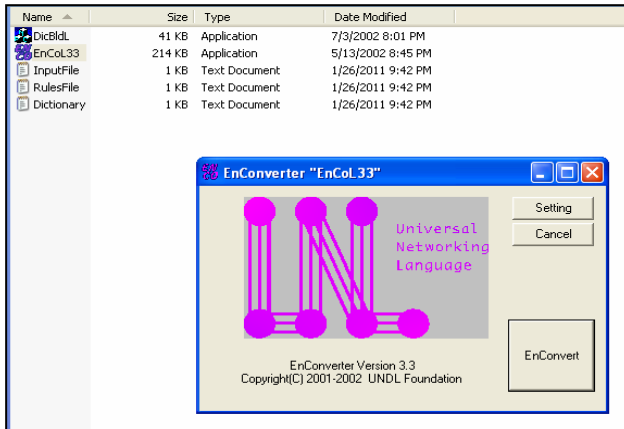**Screen 2: Creating Dictionary File**

Screen 3 shows the dictionary file *DicFile.dic* created by the *DicBldL.exe*
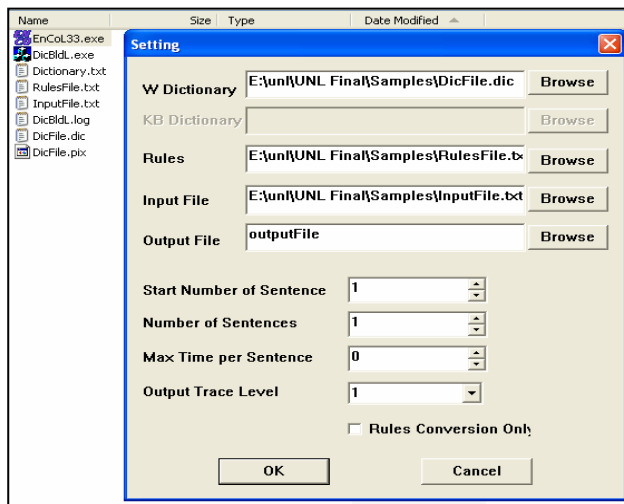

Screen 3: Dictionary File DicBldL.dic

Screen 4 shows the encoder (EnCoL33.exe) that produces UNL expression of the native language sentence
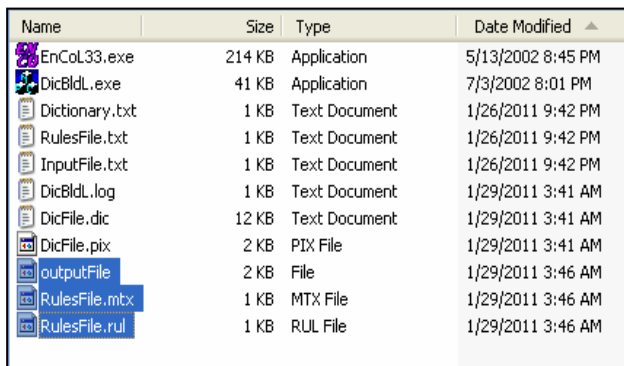
Screen 4: The encoder

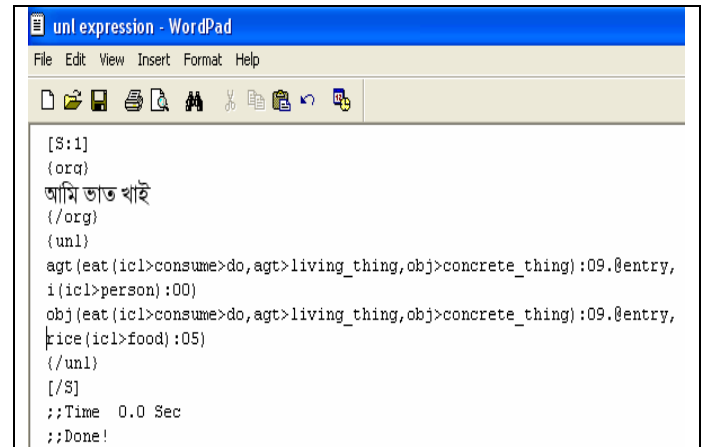Screen 5 shows the way of selecting different files to be used for conversion

Screen 5: Selecting files to be used for conversion

Screen 6 shows the output file *outputFile* created by the encoder.

Screen 6: Output file

Screen 7 shows the content of the output file i.e. UNL expression of the sentence "আমি ভাত খাই"

Screen 7: UNL Expression of the sentence "আমি ভাত খাই"

7. CONCLUSION

In this paper we have discussed all the Bangla cases and rules for converting to UNL expression. We have also demonstrated how a sentence can be converted into UNL expression by taking a sentence as an example. Although we have already worked on all the cases but for better understanding of the most of the readers of this paper we have only demonstrated one of the cases. We have mentioned this in Section 1 (Introduction). In section 2 we have outlined the methodology by which we have carried out the research. We have discussed about UNL in section 3. In section 4 and 5 we have described our work elaborately. Section 6 shows the practical implementation of our research and it is found that our rules, dictionary, etc. worked perfectly. We understand that it just a start of a long journey and hope that we would be able to reach the destination.

8. REFERENCES

[1] Md. Ershadul H. Choudhury, Md. Nawab Yousuf Ali, Mohammad Zakir Hussain Sarkar, Ahsan Razib, "Bridging Bangla to Universal Networking Language- A Human Language Neutral Meta-Language", International Conference on Computer, and Information Technology (ICCIT), Dhaka, 2005

[2] S. Abdel-Rahim, A.A. Libdeh, F. Sawalha, M. K. Odeh, "Universal Networking Language(UNL) a Means to Bridge the Digital Divide", Computer Technology Training and Industrial Studies Center, Royal Scientific Society, March 2002

[3] Uchida H., Zhu M., Della Senta T.G. "Universal Networking Language", 2005/6-UNDL Foundation, International Environment House.

[4] Uchida H., Zhu M., "The Universal Networking Language (UNL) Specification Version 3.0 1998, Technical Report, UNU, Tokyo, 1998", 2005/6-UNDL Foundation, International Environment House

[5] Enconverter Specification Version 3.3, UNU Centre, Tokyo 150-8304, Japan 2002

[6] http://www.undl.org last visited February 25, 2011

[7] D.M. Shahidullah, "Bangla Baykaron", Ahmed Mahmudul Haque of Mowla Brothers prokashani, Dhaka-2003.

[8] D. C. Shuniti Kumar, "Bhasha-Prakash Bangala Vyakaran", Rupa and Company Prokashoni, Calcutta, July 1999, pp.170-175

[9] Asad H., "Bakkotottyo", Second edition, 1994, Dhaka

[10] D. S. Rameswar, "Shadharan Vasha Biggan and Bangla Vasha", Pustok Biponi Prokashoni, November 1996, pp.358-377

[11] M.N.Y. Ali, J.K. Das, S. M. Abdullah Al Mamun, M. E. H. Choudhury, "Specific Features of a Converter of Web Documents from Bengali to Universal Networking Language", International Conference on Computer and Communication Engineering 2008 (ICCCE'08), Kuala Lumpur, Malaysia. pp. 726-731.

[12] M.N.Y. Ali, J.K. Das, S.M. Abdullah Al Mamun, A. M. Nurannabi, "Morphological Analysis of Bangla words for Universal Networking Language", International Conference on Digital Information Management, icdim, 2008, London, England, pp. 532-537

[13] M.N.Y.Ali, A. M. Nurannabi, G. F. Ahmed, J.K. Das, "Conversion of Bangla Sentence for Universal Networking Language", International Conference on Computer and Information Technology (ICCIT), Dhaka, 2010

[14] M.Z. Hossain, S. A. Noor, M. F. Mridha, "Some Proposed Standard Models for Bangla Dictionary Entries of Bangla Morphemes for Universal Networking Language", International Journal of Computer Applications (0975-8887) Volume 12- No. 6. December 2010

[15] M.N.Y. Ali, S.A.Noor, M.H.Zakir, J.K. Das, "Development of Analysis Rules for Bangla Root and Primary Suffix for Universal Networking Language", International Conference on Asian Language Processing, IALP 2010, Harbin China.

[16] http://www.undl.org/index.php?option=com_content&view=article&id=53&Itemid=99&lang=en# To download the encoder