# A Rule based Approach for Anaphora Resolution in Bengali Sentences

## Angan Das[1*], Abhishek Banerjee[2], Samiya Maity[3], Alok Ranjan Pal[4]

[1,2,3,4]Dept. of Computer Science and Engineering, College of Engineering and Management, Kolaghat, India

*Corresponding Author: angandas1195@gmail.com*

*Abstract*-Anaphora Resolution (AR) is one of the complex tasks in computational linguistics. It is the process of finding the referents in a given discourse. In this paper, anaphora resolution between a noun and its representative pronoun is presented using a rule-based model. First of all, a set of test sentences are retrieved from the Bengali text corpus. Each of the test sentences contains noun/s and its representative pronoun/s in it. Next, these sentences are processed using a set of rules to establish the relation between noun/s and its representative pronoun/s. The experiment is carried out on a set of 200 structured sentences and the overall accuracy is achieved 65.14%. The sentences are collected from the Bengali text corpus, developed in the TDIL (Technology Development for Indian Languages) project of the Government of India. The challenges and pitfalls of this work are described in this paper at last.

*Keywords*-- Bengali Anaphora Resolution, Rule based Approach, Bengali Orthographic Rule

## I. INTRODUCTION

In this report, a work is presented on Anaphora Resolution for Bengali Sentences. Anaphora resolution is one of the major challenges in Natural Language Processing. In this task, a word which refers to some earlier item in the discourse is found out. This work is on a very common appearance of anaphora resolution said pronoun resolution, where a pronoun of a simple Bengali sentence refers to a noun in the same Bengali sentence. For Example,

পুঁটি একটা বই কিনেছে কারণ সে বই পড়তে ভালোবাসে।

(Punti ektá bai kinechhe káron se bai porte válobáse)

Here, "সে" (se) is the pronoun in the second clause of the sentence and it refers to "পুঁটি" (Punti), noun in the first clause of the sentence.

However, the process of anaphora resolution becomes extremely complex for different real life sentences where anaphor creates ambiguity. Since this type of understanding is still poorly implemented in software, anaphora resolution is currently an area of active research. Introduction of anaphora resolution to many applications like machine translation, opinion mining, automatic summarization, information extraction etc. increases the accuracy of those applications to far extent

There are various works in this field in English and other European languages, but very few works are established in South Asian Languages due to its complex linguistic features. In this work, we are looking forward to propose a rule based approach to find the relation between a pronoun and a noun in a single discourse present in a Bengali sentence.

## II. SURVEY

In this section, a brief survey in this domain is presented.
One of the earliest forms of this approach was proposed by J. Hobbs [1] in 1979. According to this algorithm, one needs a parse tree or a syntax tree for a given sentence to resolve this problem. On encountering a pronoun, the parse tree of the current sentence is parsed in a left-to-right sequence until a noun is encountered with a correct gender and number agreement.

Another method was proposed by S. Lappin and H.J. Leass [2] in 1994 which used the method of salience factor. Salience is a measure which indicates how likely a noun is to co-refer to a specific pronoun or a lexical anaphora. Salience values are assigned to each referring expression. Finally, a noun with highest salience value and compatible with right pronoun agreement features is chosen as a right antecedent.
Richard Evans and Constantin Orasan [3] improved anaphora resolution by identifying animate entities in texts. In English anaphora resolution, recognition of the animacy of noun phrases (NP) improves the accuracy with which gender agreement restrictions can be enforced between pronouns and candidates. The work illustrates the animacy of English NPs which is identified using a combination of a number of tactics.

A. Dhar and U. Garain [4] attempted to find an optimized set of salience values for doing anaphora resolution in Bengali. This paper attempts to use an off-the-shelf anaphora

resolution (AR) system for Bengali. The language specific pre-processing modules of GuiTAR (v3.0.3) are identified and suitably designed for Bengali. Anaphora resolution module is also modified or replaced in order to realize different configurations of GuiTAR. Performance of each configuration is evaluated and experiment shows that the off-the-shelf AR system can be effectively used for Indic languages.

In the work of T. Tazakka, Md. Asifuzzaman and Sabir Ismail [5], pronouns and verbs are considered as anaphor and nouns are considered as its antecedent. It shows that, anaphor and antecedent in Bengali matches with each other in some factors like Number, Gender and Person. In case of personal pronouns in Bangla, status of person can be differentiated such as Honorable, Normal and Negotiable. Experimenting with these factors proved that this method works well and give accuracy of around 80%.

The work proposed by Anwesa Bagchi [6] focuses on the concept of origin and functional behavior of Postpositions in Bangla with special reference to Prepositions in English. It has been often observed that Postpositions in Bangla function as case endings that develop syntactic relationship between different syntactic units. Furthermore, Postpositions carry valuable semantic information about the relationship between the noun and the verb. This study concentrates on these aspects of the Postpositions in Bengali with special reference to similar functions performed by Prepositions in English.

In the work proposed by U.K. Sikdar and group [7] demonstrates adapting a state-of-the-art anaphora resolution system for a resource poor language like Bengali. Performance of any anaphoric resolver greatly depends on the quality of a mention detector. A number of models for mention detection have been developed based on heuristics and machine learning. The experiments show that a language-dependent system can attain reasonably good performance when re-trained on a new language with a proper subset of features.

### III.　PRE-PROCESSING

In this work, first of all, a series of text pre-processing tasks are executed, as-

*A. Text Normalization*

Before the experiment gets started, texts are normalized through a series of text normalization steps, like- a) detachment of punctuation marks like single quote ('), tilde (~), double quote (""), parenthesis (), comma (,) etc.; b) conversion of dissimilar fonts into similar one; c) removal of angular brackets (<>), uneven space, broken lines, slashes etc. from sentences; d) identification of sentences terminal

markers (i.e., dāri ("।"), note of exclamation, and note of interrogation), etc.)

*B. A Sample non-Normalized Text*

A text, retrieved from the Bengali text corpus is non-normalized in nature (refer Figure 1).

সংখ্যাতথ্য বলছে , মহিলাদের মধ্যে ১৯৮১ খৃষ্টাব্দে সুমারীঅনুযায়ী সাক্ষরতার হার মালদা , মুর্শিদাবাদ , পুরুলিয়া , পশ্চিম দিনাজপুর এবংকোচবিহার জেলায় যথাক্রমে শতকরা হিসাবে ১৪.২১ , ১৭.২৮ , ১৩.৭৪ , ১৭.০৮ , এবং ১৯.০৭।

Figure 1. Sample non-normalized text

*C. A Sample Text aftert Normalization*

A sample normalized text, generated after the normalization process is given below (refer Figure 2).

সংখ্যাতথ্য বলছে মহিলাদের মধ্যে ১৯৮১ খৃষ্টাব্দে সুমারীঅনুযায়ী সাক্ষরতার হার মালদা মুর্শিদাবাদ পুরুলিয়া পশ্চিম দিনাজপুর এবংকোচবিহার জেলায় যথাক্রমে শতকরা হিসাবে ১৪২১ ১৭২৮ ১৩৭৪ ১৭০৮ এবং ১৯০৭।

Figure 2. Sample normalized text

*D. Text Lemmatization*

In every text processing related work, Lemmatization is one of the most important pre-processing tasks to increase the lexical coverage of the data. In this experiment, the stem part and the POS of every word are evaluated manually.
A sample lemmatized text after manual text lemmatization is presented below-

সংখ্যাতথ্য/সংখ্যাতথ্য/n বলছে/বলা/vb মহিলাদের/মহিলা/n মধ্যে/মধ্যে/adj ১৯৮১/১৯৮১/n খৃষ্টাব্দে/খৃস্টাব্দ/n সুমারী/ সুমারী/n অনুযায়ী/অনুযায়ী/adj সাক্ষরতার/সাক্ষর/n হার/হার/vb মালদা/মালদা/n মুর্শিদাবাদ/মুর্শিদাবাদ/n পুরুলিয়া/পুরুলিয়া/n পশ্চিম/পশ্চিম/n দিনাজপুর/দিনাজপুর/n এবং/এবং/conj কোচবিহার/কোচবিহার/n জেলায়/জেলা/n যথাক্রমে/যথাক্রম/adj শতকরা/শতকরা/adv হিসাবে/হিসাব/adj ১৪২১/১৪২১/n ১৭২৮/১৭২৮/n ১৩৭৪/১৩৭৪/n ১৭০৮/১৭০৮/n এবং/এবং/conj ১৯০৭/১৯০৭/n ।

Figure 3. Sample lemmatized text

*E. Identification of Conjunction*

In this work, the relation between the noun and its representative pronoun is resolved based on an assumption that the noun and its representative pronoun are present in two different clauses. So, the first task was to identify the conjunction present in the sentence. Considering the

　　　　　　　　　　　　　　　　　　　**30**

conjunction of the sentence, two clauses are obtained. For Example, in this sentence-

পুঁটি গল্পের বই পড়ছে **কারণ** সে বই পড়তে ভালোবাসে।

Figure 4.Identifiation of conjunction

the word "এবং" is a conjunction. So, this word is identified by the tagged Part-of-Speech of this word.

## IV.    PROPOSED APPROACH

In the proposed approach, first the target sentences are collected from the corpus. Next, the sentences are passed through the text normalization (refer section *3.A*) and lemmatization (refer section 3.*B*) processes.
After that, the whole sentence is scanned for the pronoun/s. In the proposed work, pronouns are categorized into 4 groups:
a)      Singular Person Pronoun
b)      Singular Object Pronoun
c)      Plural Person Pronoun
d)      Plural Object Pronoun
Next, by identifying the conjunction word, the sentence is divided into two separate clauses.
After that, the noun/s in the sentence is/are also identified. The nouns are categorized into two groups:
a)      Person Noun
b)      Object Noun
Finally, these nouns and pronouns are considered for evaluation and a relation is established between them based on some case (or rule).

The cases are as follows:
**Case 1: One Person Noun and One Person Pronoun**
If the first clause contains one person noun, and the second clause contains one person pronoun, then the pronoun definitely points to the noun.
For Example: পুঁটি গল্পের বই পড়ছে কারণ সে বই পড়তে ভালোবাসে।
(Punti galper bai parche káron se bai parte válobáse)
Here, "পুঁটি" (Punti) singular person noun and "সে" (se) is singular person pronoun. So, the pronoun "সে" (se) means "পুঁটি" (Punti).

**Case 2: One Object Noun and One Object Pronoun**
If the first clause contains one object noun, and the second clause contains one object pronoun, then the pronoun definitely points to the noun.
For Example: কাগজটি ফেলে দেওয়াই ভাল কারন এটি আর কোন কাজে লাগবে না।
(Kágajti fele deoyái válo káran eti ár kona káje lágbe ná)

Here,"কাগজটি" (Kágajti) is a singular object noun and "এটি" (eti) is a singular object pronoun. So, the pronoun "এটি" (eti) definitely indicates the noun "কাগজটি" (Kágajti).

**Case 3: One Person Noun, One Object Noun and One Person Pronoun**
If the first clause contains one person noun and one object noun, and the second clause contains only a person pronoun, then the person pronoun represents the person noun only.
For Example: রাম আজ স্কুল যায়নি কারন সে অসুস্থ ছিল।
(Rám áj skul jáyni káran se asustha chhila)
Here, "রাম" (Rám) is a person noun and "স্কুল" (Skul) is an object noun and "সে" (se) is a singular person pronoun. Obviously, the person pronoun "সে" (se) represents the person noun "রাম" (Rám).

**Case 4: One Person Noun, One Object Noun and One Object Pronoun**
If the first clause contains one person noun and one object noun, and the second clause contains only an object pronoun, then the object pronoun represents the object noun itself.
For Example: কৃতিকা অঙ্কন প্রতিযোগীতার প্রাইজটা পেল ঠিকই কিন্তু সেটা রাজেরই প্রাপ্য ছিল।
(Krittiká ankan pratijyogitár praijtá pela thiki kintu setá Rájeri prápya)
Here, "কৃতিকা" (Krittiká) is a person noun and "প্রাইজ" (práij) is an object noun and "সেটা" (setá) is a singular object pronoun. Obviously, the object pronoun "সেটা" (setá) refers to the object noun "প্রাইজটা" (práijtá).

**Case 5: Two Person Nouns and One Compound Pronoun**
If the first clause contains two person nouns and the second clause contains one compound pronoun, then the pronoun definitely points to those two nouns in the first clause.
For Example: রাম এবং লক্ষন বনবাসে গেছিল ঠিকই কিন্তু তাদের যায়গা আর কেউ নিতেপারেনি কখনই।
(Rám eban Lakshan banabáse gechhila thiki kintu táder jáygá ár keu nite páreni kakhanai)
Here, "রাম" (Rám) and "লক্ষন" (Lakshan) are person nouns and "তাদের" (táder) is a compound pronoun. Obviously, the compound pronoun indicates the nouns "রাম" (Rám) and "লক্ষন" (Lakshan).

**Case 6: Two Object Nouns and One Compound Pronoun**
If the first clause contains two object nouns and the second clause contains one compound pronoun, then the pronoun definitely points to those two nouns in the first clause.
For Example: পুরনো টেবিল আর চেয়ার গুলো ফেলে দেওয়াই ভাল কারন ওগুলোতে উই ধরে গেছে।

(Purono tebil ár cheyár gulo fele deoyái válo)
Here, "টেবিল" (tebil) and "চেয়ার" (cheyár) are object nouns and "ওগুলোতে" (ogulote) is a compound pronoun. Obviously, the compound pronoun indicates the nouns "টেবিল" (tebil) and "চেয়ার" (cheyár).

### Case 7: One Person Noun, One Object Noun, One Person Pronoun and One Object Pronoun

If the first clause contains one person noun and one object pronoun and, similarly, the second clause contains one person pronoun and one object pronoun, then the person pronoun points to the person noun in the first clause and in the same way, the object pronoun points to the object noun.

For Example: নীহারিকার কাছে গুচ্ছ বই আছে কিন্তু সে সেগুলোর গুরুত্ব বোঝে না।

(Nihárikár káche guchchha bai áchhe kintu se segulor gurutta bojhe ná)

Here, "নীহারিকার" (Neehárikár) is a person noun, "বই" (bai) is an object noun, "সে" (se) is a singular person pronoun and "সেগুলোর" (segulor) is an object pronoun. Obviously, the person pronoun "সে" (se) indicates the person "নীহারিকার" (Nihárikár) and object pronoun "সেগুলোর" (segulor) represents "বই" (bai).

### Flowchart



```
        Start

Sample Bengali text from Bengali
         text corpus

     Texts are Normalized
   Normalized texts obtained

     Texts are Lemmatized
    Lemmatized text obtained

 Sentences are split into two clauses

    With the help of some cases
 Relation between nouns and its
 corresponding pronouns are
 established with the help of some
 rules

          END
```
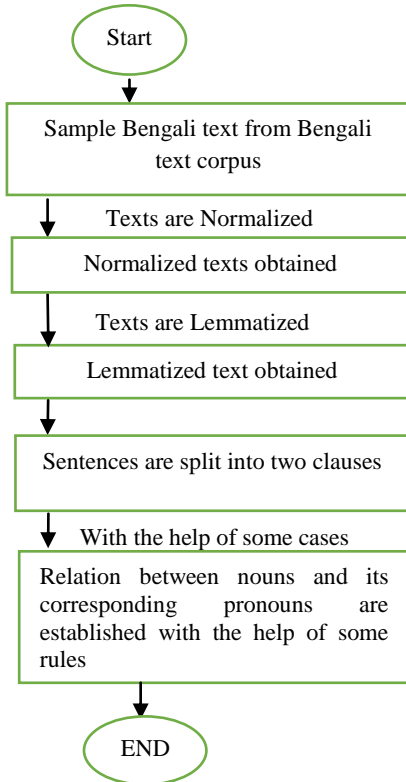
Figure 5. Flowchart of the overall approach

### Legends Used in the Algorithm

To determine the category of the nouns and pronouns the following legends have been used:

| Noun Category | POS tag used |
|---|---|
| 1) Person Noun | pn |
| 2) Object Noun | on |

| Pronoun Category | POS Tag Used |
|---|---|
| 1) Person Pronoun Singular | pprns |
| 2) Object Pronoun Singular | oprns |
| 3) Person Pronoun Plural | pprnp |
| 4) Object Pronoun Plural | oprnp |

| Parts of Speech | POS Tag Used |
|---|---|
| 1) Conjunctions | conj |
| 2) Others | pos |

## V. OUTPUT AND CORRESPONDING EVALUATION

By applying the above mentioned cases, the following results have been obtained:

### Case 1: Sentence containing one Person Noun and one Person Pronoun.

Input:
রাহিমের/রাহিম/pn     মতে/মত/pos     পুলিশ/পুলিশ/pos
বর্তমান/বর্তমান/pos     হত্যাকাণ্ড/হত্যাকাণ্ড/pos     সম্পর্কে/সম্পর্ক/pos
যথেষ্ট/যথেষ্ট/pos     তৎপর/তৎপর/pos     তবুও/তবু/conj
তার/তার/pprns     আশঙ্কা/আশঙ্কা/pos     এই/এই/pos
তৎপরতা/তৎপর/pos     শেষ/শেষ/pos     পর্যন্ত/পর্যন্ত/pos
থাকবে/থাকা/pos কিনা/কিনা/pos ।

Output:
In this sentence, "রহিম" is the Person Noun in the first clause and "তার" is the Person Pronoun in the second clause. So, after evaluating the sentence by the proposed algorithm, the result is obtained as:

**"তার" represents the noun word "রহিম"**

### Case 2: Sentence containing one Object Noun and one Object Pronoun.

Input:
গতকাল/গতকাল/pos চুরি/চুরি/pos হওয়া/হওয়া/pos চিঠিটি/চিঠ/on
পাওয়া/পাওয়া/pos     গিয়েছে/যাওয়া/pos     কিন্তু/কিন্তু/conj
এটি/এটি/oprns আদেও/আদেও/pos আসল/আসল/pos কিনা/কিনা/pos
এখনো/এখন/pos জানা/জানা/pos যায়নি/যাওয়া/pos ।

Output:
In the above sentence, "চিঠিটি" is an Object Noun (present in the first clause) and "এটি" is the Singular Object Pronoun

(present in the second clause). So, after passing the sentence through the algorithm, we get the output as:

**"এটি" represents the noun word "খবর"**

### Case 3: Sentence containing one Person Noun, one Object Noun and one Person Pronoun.

Input:

সামারস্পেশালে/সামারস্পেশাল/on পঞ্চননবাবুর/পঞ্চননবাবু/pn রিজারভেশন/রিজারভেশন/pos ছিল/ছিল/pos কিন্তু/কিন্তু/conj ওইদিন/ওইদিন/pos দুপুরে/দুপুর/pos ট্রেন/ট্রেন/pos ধরতে/ধরা/pos স্টেশনে/স্টেশন/pos এসে/আসা/pos তিনি/তিনি/pprns দেখেন/দেখা/pos রিজারভেশন/রিজারভেশন/pos চারটে/চার্ট/pos নাম/নাম/pos নেই/নেই/pos ।

Output:

In the above sentence, "পঞ্চননবাবু" is a Person Noun (present in the first clause) and "সামারস্পেশাল" is an Object Noun (present in the first clause) and "তিনি" is the Singular Person Pronoun (present in the second clause). So, after passing the sentence through the algorithm, we get the output as:

**"তিনি" represents the noun word "পঞ্চননবাবু"**

### Case 4: Sentence containing one Person Noun, one Object Noun and one Object Pronoun.

Input:

চাওলা/চাওলা/pn শহরে/শহর/pos এসে/আসা/pos একটি/এক/pos হোটেলে/হোটেল/on থাকতে/থাকা/pos চেয়েছিলেন/চাওয়া/pos কিন্তু/কিন্তু/conj তার/তার/oprns বাইরে/বাইরে/pos দুপক্ষে/দুপক্ষে/pos রীতিমত/রীতিমত/pos হাতাহাতি/হাতাহাতি/pos হয়ে/হওয়া/pos গেছে/যাওয়া/pos কয়েকবার/কয়েকবার/pos ।

Output:

In the above sentence, "চাওলা" is a Person Noun (present in the first clause) and "হোটেল" is an Object Noun (present in the first clause) and "তার" is the Singular Person Pronoun (present in the second clause). So, after passing the sentence through the algorithm, we get the output as:

**"তার" represents the noun word "হোটেল"**

### Case 5: Sentence containing Two Person Nouns and one Compound Pronoun.

Input:

গোরা/গোরা/pn ৩/৩/pos বিনয়/বিনয়/pn অভিন্ন/অভিন্ন/pos হৃদয়/হৃদয়/pos বন্ধু/বন্ধু/pos হলেও/হলেও/conj তাদের/তারা/pprnp স্বভাব/স্বভাব/pos চরিত্র/চরিত্র/pos একেবারে/একবার/pos একরকম/একরকম/pos ছিল/ছিল/pos না/না/pos ।

Output:

In this sentence, "গোরা" and "বিনয়" are the Person Nouns in the first clause and "তাদের" is the Compound Pronoun in the

second clause. So, after passing the sentence through the proposed algorithm, we get the result as:

**"তাদের" represents the noun word "গোরা" and "বিনয়"**

### Case 6: Sentence containing Two Object Nouns and one Compound Pronoun.

Input:

শিমুল/শিমুল/on ৩/৩/pos পলাশ/পলাশ/on এমন/এমন/pos দুটি/দুটি/pos গাছ/গাছ/pos যে/যে/pos সারাবছর/সারাবছর/pos কারও/কার/pos চোখে/চোখ/pos পড়ে/পড়া/pos না/না/pos কিন্তু/কিন্তু/conj বসন্তকালে/বসন্তকাল/pos এরা/এরা/oprnp পথিকের/পথিক/pos দৃষ্টি/দৃষ্টি/pos আকর্ষণ/আকর্ষণ/pos করেই/করা/pos ।

Output:

In this sentence, "শিমুল" and "পলাশ" are the Object Nouns in the first clause and "এরা" is the Compound Pronoun in the second clause. So, after passing the sentence through the proposed algorithm, we get the result as:

**"এরা" represents the noun word "শিমুল" and "পলাশ"**

### Case 7: Sentence containing one Person Noun, one Object Noun, one Person Pronoun and one Object Pronoun.

Input:

পর্যটকেরা/পর্যটক/pn কলকাতায়/কলকাতা/on আসতে/আসা/pos চান/চাওয়া/pos কিন্তু/কিন্তু/conj এথানে/এথানে/oprns ভাল/ভাল/pos হোটেলের/হোটেল/pos ব্যবস্থা/ব্যবস্থা/pos কম/কম/pos বলে/বলা/pos তাঁরা/তাঁরা/pprns আসেননা/আসা/pos ।

Output:

In this sentence, "পর্যটক" and "কলকাতা" are the Person Noun and Object Noun respectively in the first clause and "তারা" and "এথানে" are the Singular Person Pronoun and Singular Object Pronoun, respectively, in the second clause. So, after passing the sentence through the proposed algorithm, we get the result as:

**"তারা" represents the noun word "পর্যটক"**

In the same way, all the seven cases are implemented on total 200 sentences and the percentage of accuracy is obtained as given below-

Table 1. Overall Performance of the System

| Cases | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-------|---|---|---|---|---|---|---|
| **Accuracy** | 71 | 76 | 65 | 59 | 62 | 63 | 60 |

Therefore, the overall accuracy is achieved as 65.14%.

## VI.    FEW CLOSE OBSERVATION

Some of the challenges which are observed closely during the experiment are given below:

a. As there is not any tool for all word lemmatization in Bengali, so the lemmatization task is performed manually.

b. As the real life sentences are very much complex in nature, they can't be framed using any predefined rule. It's the main challenge in this work.

c. Person Names in conjugate form (like "রবিন্দ্রনাথ ঠাকুর", "নরেন্দ্র দামোদরদাস মোদী" etc.) require separate rule for processing.

## VII.   CONCLUSION AND FUTURE WORK

This work establishes the relation between a noun and its corresponding pronoun on structured sentences using a rule-based system. Various works have been done on English language but the works on resource poor languages like Asian languages are comparatively less in number.

Some future works are yet to be done, such as-

a. Semantically related sentences are yet not covered in the approach.

b. Simple sentences which do not contain any conjunction but contains a noun and its corresponding pronoun in a single sentence have not been considered in this model.

c. Complex sentences containing more than two clauses have not been considered as well.

## REFERENCE

[1] Hobbs, Jerry, "Resolving pronoun references", In the Proceedings of the Lingua 44:311-338, 1979.

[2] S. Lappin and H.J. Leass, "An algorithm for pronominal anaphora resolution", In the Proceedings of the Computational Linguistics, 20(4), 535-561, 1994.

[3] Constantin Orasan and Richard Evans,"NP Animacy Identification for Anaphora Resolution", Journal of Artificial Intelligence Research 29, 79-103, 2007.

[4] A. Dhar and U. Garain, "A method for pronominal anaphora resolution in Bengali", In the Proceedings of the 6th Int. Conf. on Natural Language Processing (ICON), Student Competition section, Pune, December.

[5] T Tazakka, Md. Asifuzzaman and Sabir Ismail, "Anaphora Resolution in Bangla Language", International Journal of Computer Applications (0975 – 8887), November 2016.

[6] Anwesa Bagchi, "POSTPOSITIONS IN BANGLA With special reference to Prepositions in English", LANGUAGE IN INDIA, Volume 7 : 11 November, ISSN 1930-2940, 2007.

[7] U.K. Sikdar, A. Ekbal, S. Saha, Olga Uryupina and M. Poesio, "Adapting a state-of-the-art Anaphora Resolution System for Resource-poor Language", In the Proceedings of the International Joint Conference on Natural Language Processing, pages 815–821, Nagoya, Japan, October 2013.

**Authors Profile**

*Mr. Angan Das, Mr. Abhishek Banerjee* and *Miss. Samiya Maity* are the students of final year in Computer Science and Engineering Dept., College of Engineering and Management, Kolaghat.

*Dr. Alok Ranjan Pal* is an assistant professor in the same dept. since 2006. His research area is Computational Linguistics.