

# Prompt-Based Anaphora Resolution in Large Language Models

Mridusmita Das\*, Apurbalal Senapati

Central Institute of Technology Kokrajhar,  
Department of Computer Science and Engineering,  
India

{ph21cse1003, a.senapati}@cit.ac.in

**Abstract.** With the advancement of Large Language Models (LLMs), the scope of research in the Natural Language Processing (NLP) domain has significantly shifted. The LLM has context-based advanced language understanding that is suitable for various types of discourse analysis. Creating suitable prompts can effectively guide the model's responses toward the desired outcome. Anaphora resolution is a complex problem that is highly context-dependent. This paper attempted to explore a prompt-based LLM technique for the resolution of anaphora. Our experiment used a **text-based question** prompt within the OpenAI LLM framework. The experiment is conducted in the Assamese language, initially using a rule-based system. The results are then compared with those obtained from a prompt-based approach. The main contribution of this paper is the exploration of prompt engineering techniques for anaphora resolution. The results indicate that the prompt-based approach is significantly superior to the rule-based approach.

**Keywords.** Anaphora resolution, prompt engineering, large language models, Assamese language, low-resource NLP, peer-to-peer network.

## 1 Introduction

Anaphora and coreference resolution are a longstanding and classical problem in both linguistic and computational aspects of the NLP domain. Research in this area began decades ago and remains an active field, particularly for low-resource languages. **Anaphora** is defined as a linguistic entity that refers to other entities mentioned earlier in the text [11]. On the other hand, **coreference** refers to two or more words or

phrases in a text that refer to the same entity [5]. Therefore, **anaphora** is a subset of **coreference**.

*Students are passionate about NLP. They spend hours every day learning it. The subject is very demanding today.*

Consider the above text, where the anaphora **They** refers to **Students** earlier in the text. On the other hand, **NLP** and **The subject** are coreferential, as both refer to the same entity in the context. In this paper, we focus exclusively on anaphora resolution.

Effective resolution of the anaphora is crucial for accurate language understanding. Anaphora resolution plays an essential part in many NLP applications, such as question answering, text summarisation, and machine translation [10]. Most anaphora resolution systems focus on resolving anaphors whose antecedents are noun phrases. However, antecedents can also be verb phrases, clauses, sentences, paragraphs, or even larger discourse segments [8]. In many cases, resolving this requires not only grammatical and contextual information but also real-world knowledge [14]. Nowadays, researchers are increasingly focusing on more complex cases of anaphora [27]. The pronominal anaphora refers to a specific type of anaphora i.e., a pronoun (like he, she, it, they, etc.) refers back to a noun phrase or entity that was mentioned earlier in the text [10]. As a pioneering attempt at resolution of anaphora through prompt engineering, this paper focuses solely on the resolution of pronominal anaphora in the Assamese language.

Accurate resolution of anaphora is a critical requirement for downstream NLP applications such as information extraction, machine translation, question answering, and text summarisation. Traditional systems have largely concentrated on resolving noun phrase antecedents, although in practice, antecedents may also be verb phrases, clauses, or larger discourse segments. Although substantial progress has been achieved in English and a few other languages of high resources, research on Indian languages, including Assamese, remains limited.

Assamese poses unique challenges due to its agglutinative morphology, case-marking system, and relatively free word order. Moreover, the scarcity of annotated corpora and computational resources makes the task of anaphora resolution more complex in this low-resource setting. To date, most existing work on Assamese has explored rule-based or statistical approaches, but little attention has been paid to leveraging LLMs. Recent advances in LLMs have demonstrated strong capabilities in discourse understanding, which has prompted researchers to investigate whether prompt engineering can guide these models toward more accurate resolution of context-dependent phenomena such as anaphora.

However, the potential of prompt-based approaches for the Assamese language remains underexplored. Given the linguistic complexity of Assamese and the absence of prior systematic studies on prompt-based anaphora resolution, this research investigates whether carefully designed prompts for LLMs can outperform traditional rule-based techniques in resolving pronominal anaphora in Assamese texts.

The major contributions of this paper are summarized as follows:

**Corpus Development:** We constructed a domain-diverse Assamese text corpus, curated through web scraping and preprocessing, containing rich instances of pronominal anaphora.

**Rule-Based Baseline:** We implemented a peer-to-peer network-based rule-driven approach as a baseline for Assamese anaphora resolution.

**Prompt-Based Technique:** We proposed and evaluated a text-based question prompt method

with LLMs for resolving Assamese pronominal anaphora.

**Comparative Evaluation:** We present a comparative study of rule-based and prompt-based approaches, showing that prompt engineering yields superior performance and sets a benchmark for future studies.

This pioneering attempt to apply prompt engineering for pronominal anaphora resolution in the Assamese language aims to not only advance computational studies in this under-represented language but also provide a foundation for further research in discourse processing for low-resource settings.

## 2 Related Work

The earlier research on anaphora resolution was a rule-based approach, relying on various linguistic features. These rules were manually crafted by observing specific patterns in the text. The rules are primarily based on grammatical features such as gender agreement, number agreement, c-command constraints, syntactic parallelism, semantic consistency, semantic parallelism, salience, proximity, etc. [7, 10, 19]. Mitkov conducted extensive work, analyzed various aspects, and identified the state of the art at that time [10, 11]. He primarily focused on the rule-based approach.

A major limitation of the rule-based approach is its reliance on annotated datasets. Rule-based algorithms cannot be directly applied to raw text. They require a certain level of annotation, such as part-of-speech (POS) tags, named entity (NE) tags, chunking information, parsing, and morphological details [16]. In late 1990, research in anaphora resolution accelerated because of the availability of tagged data and preprocessing tools such as POS taggers, NE taggers, and parsers. Around 2000, machine learning approaches for anaphora resolution were introduced. Initially, statistical and expectation maximization techniques were used by incorporating existing grammatical features into the probabilistic model [1, 6]. Soon et al. [23] used tagged coreference chains in a machine learning approach, where a statistical, off-the-shelf anaphora resolution system based on

BERT [24] has emerged. Gradually, researchers began to address more complex features of anaphora [13, 18].

One of the pioneering attempts using Deep Learning (DL) was made by Manning et al. [2]. Based on the words in the mentions and their context, they used neural networks to automatically create dense vector representations for mention pairs. Subsequently, they employed reinforcement learning to further optimize the system [3]. Various results were presented in the CoNLL-2012 shared task [17] for predicting coreference in Chinese, English, and Arabic. Wiseman et al. [25] used recurrent neural networks (RNNs) to model mentions. More recently, Plu et al. [15] attempted to improve the Stanford 'deep-coref' system. Additionally, Yang et al. [26] used LLMs to explore how ChatGPT-like models perform on anaphora resolution.

Compared to English-like languages, work on Indian languages is limited [22]. Pakray et al. [12] incorporated a lexical, syntactic, and semantic framework into a textual entailment recognition system. However, many researchers have made efforts in languages such as Telugu, Bengali, Malayalam, Hindi, Assamese, etc. Sobha et al. [9] tried to include a generic model, created a resolution mechanism, and handled a range of difficulties in several languages. Senapati et al. [20] tried to adapt the English-based system for Indian languages with the necessary modifications. Sikdar et al. [21] used a feature selection method based on differential evolution (DE) to resolve anaphoras in Bengali, a language with limited resources. Das et al. [5] proposed a novel approach to identifying coreference chains using the peer-to-peer network concept. They also attempted to resolve coreference using basic prompt techniques such as zero-shot, few-shot and CoT prompting in Assamese [4].

### 3 System Description

The task is divided into four primary stages: corpus development, anaphora resolution in a rule-based approach, anaphora resolution in a prompt-based approach, and results with error analysis. Section 3.1 provides specifics on the

data collection procedure. The second phase, anaphora resolution in a rule-based approach, is the most crucial step for extracting pronominal anaphora and their contexts from the corpus. The objective of the third phase is to design a question-based prompt that yields the desired outcome, whereas the fourth phase is dedicated to analyzing the results.

#### 3.1 Corpus Development

We have taken into consideration two criteria for choosing texts: the texts must be from a broad topic and have a sufficient number of pronominal anaphora. We have created a web scraper that retrieves the corpus using the Beautiful Soup<sup>1</sup> Python package, saving each item as raw text in a *.txt* file.

A URL filtering technique has been implemented in order to retrieve news from a certain domain. In order to get rid of undesirable components, including HTML tags, numbers, hyperlinks, ad text, photos, and special characters, the cleaning procedure is carried out as a pretreatment step. Next, UTF-8 encoding is used to store cleaned texts.

Table 1 shows the corpus volume. Ten distinct areas are included in the dataset, including Assamese novel text<sup>2</sup>, biopic text<sup>3</sup>, news text<sup>4</sup>, short stories<sup>5</sup>, sports<sup>6</sup>, study material text<sup>7</sup>, tourism text<sup>8</sup>, and so on. For the purposes of this particular work, we combined the stories into a single, substantial corpus. There are about 4,678 tokens and 752 lines in the collection.

<sup>1</sup>[https://www.crummy.com/software/BeautifulSoup/?utm\\_source=chatgpt.com](https://www.crummy.com/software/BeautifulSoup/?utm_source=chatgpt.com)

<sup>2</sup><https://www.newjobsinassam.com/gk/download-best-assamese-books-pdf/>

<sup>3</sup><https://assamintro.com/mamoni-raison-goswami-biography-in-assamese/>

<sup>4</sup><https://dainikjanambhumi.co.in/>

<sup>5</sup><https://www.jonakaxom.in/2019/08/assamese-short-story.html>

<sup>6</sup><https://as.wikipedia.org/sports>

<sup>7</sup><https://dte.assam.gov.in/portlets/study-materials>

<sup>8</sup><https://as.wikipedia.org/wiki/tourism>

**Table 1.** Volume of the corpus

S. No.	Description	Size
1	Number of Stories	10
2	Number of Sentences	752
3	Number of tokens	4678

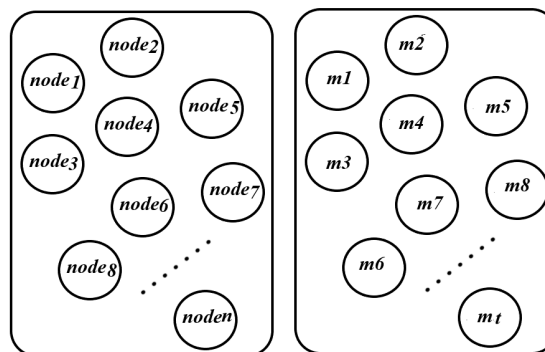
### 3.2 Pronominal Anaphora Resolution

The pronoun is an important parts-of-speech in any language. By definition, pronouns are the linguistic expressions that are used instead of nouns or noun phrases in the earlier text. When there is a use of the pronoun in a text or in speech the readers or listeners easily know which nouns you are referring to. But this pronoun resolution is not an easy task for a machine, and it is now a challenging task in language processing. On the other hand, due to the agglutinative nature of the Indic languages, pronouns are more complex compared to English-like languages.

Hence it is an essential task to study pronouns extensively in order to build sophisticated language processing applications. In this context, this paper focuses on a corpus-based study of Assamese pronominal prompt-based anaphora resolution. Based on our knowledge, there is no such corpus-based study of Assamese pronominal anaphora using prompt techniques present. The study not only focuses on the linguistic aspects but also focuses on the descriptive statistics.

For this study, we have used the Assamese corpus mentioned in section 3.1. The Assamese pronouns are inflected only on the suffixes. The language is a Neo-Indo-Aryan language that has been recognized by the Indian constitution as one of the official languages spoken in Assam.

As of the 2011 Census, Assamese had approximately 15.1 million native speakers and over 22 million total speakers (including second-language users). Assamese is an agglutinative language with an SOV word order. It is a language that uses both nominative and accusative pronouns.

**Fig. 1.** Nodes in a Network vs Mentions in a text

#### 3.2.1 Peer-to-peer: A Rule-based Approach for Anaphora Resolution

It is a relatively new approach proposed by Das et al. [5] based on networking concepts, where each mention or antecedent is represented as a node. These nodes are connected to one another according to semantic and grammatical criteria. Since all nodes exist at the same level i.e., without any hierarchy or client-server relationship and hence, the approach is referred to as peer-to-peer.

The concept is illustrated in Fig. 1. On the left side of the figure represents network nodes that are intended to be linked with similar nodes based on certain agreements. On the right side of the figure, it shows the mentions (nouns, pronouns, and noun phrases) identified in a text. These mentions also seek to connect or link with one another according to grammatical and semantic agreements.

The system was originally developed for coreference resolution, but in this work, we have adapted it for anaphora resolution. Most anaphora resolution techniques try to resolve the anaphora relation either locally or at the sentence level. Here, we use the peer-to-peer network for a coreference resolution system [5] to resolve the pronominal anaphora of the entire dataset.

It attempts to resolve its backward direction in the same sentence or previous sentences when it encounters an anaphora. In contrast, all of the nouns and pronouns are regarded as distinct entities across the whole text (Fig. 2). All of an entity's syntactic and semantic properties are

included (Fig. 3). Here, the mentions with similar colors represent semantic similarity. After that, the entities are divided according to their semantic data. In this case, the semantic information suggests the main categories of entities, such as a person, a place, etc. In an attempt to create a network or linkages, the other entities try to connect with other nodes. Nodes are joined in this manner, and the network progressively expands (Fig. 4). But here we are going to resolve only pronominal anaphora, not coreference links.

So, we can use this rule-based system [5] for getting the anaphors from the antecedents. Within a peer-to-peer network context, the resolution process is a rule-based method predicated on certain lexical and grammatical agreements. The grammatical information comprises their semantic categories, such as animate, inanimate, person, location, institution, etc. (the criterion of partitioning). Conversely, grammatical information comprises honorific information, person, number, and gender. Lexical information is given, including the sentence number (where it belongs), token number (token number in the sentence), and so on.

All of the entities in a given partition are aliases to one another and are regarded as a peer-to-peer network. When a node wants to join a peer-to-peer network or when an external entity wants to refer to any one entity, any node in the network is allowed to become a new node. Here, we have introduced the concept of a virtual node, which acts as the gateway to a network. This implies that a new entity or pronoun will submit a request to the virtual node to refer any node, and the node will decide whether to accept or reject it based on grammatical and other agreements.

In the formation of a network or links among nodes, there exists the concept of a virtual node. A virtual node represents the combined lexical, grammatical, and semantic information of all nodes within the network. This concept arises only when a network or link is established with two or more nodes. In the case of an individual node, the node itself acts as the virtual node. When a new node seeks to join the network, it interacts with the virtual node, which based on predefined agreements (rules) either permits or

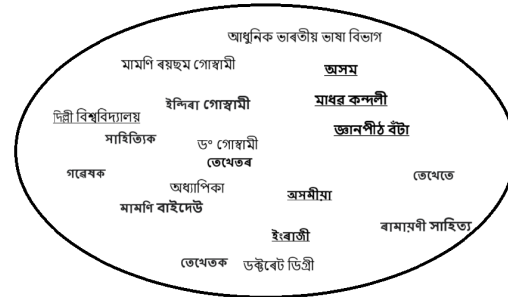


Fig. 2. Initial Cluster with all Mentions



Fig. 3. Categorized them Semantically (Same colour represent semantic similarity)

rejects the connection. The technique of virtual node generation is illustrated in Fig. 5.

Initially, all the mentions are partitioned based on semantic similarity, as illustrated with different colours in the Fig. 3. A virtual node with all of the grammatical and lexical information for every partition is created because all of the entities are aliases. The two nodes ( $n_1$  and  $n_2$ ) and their grammatical and lexical information (attributes) are shown. The new virtual node (Fig. 5), representing nodes ( $n_1$ ) and ( $n_2$ ), has the attribute defined as:  $Attribute(Virtual\ node) = Attribute(n_1) \cup Attribute(n_2)$ .

When an external entity or pronoun tries to link with any entities of a partition, or when any external entities or pronouns try to participate in a partition, the virtual node will decide based on the agreements and regulations.

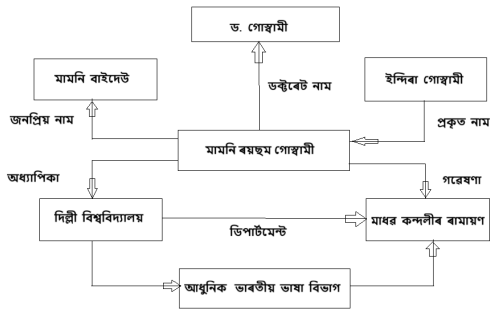


Fig. 4. Mention Links

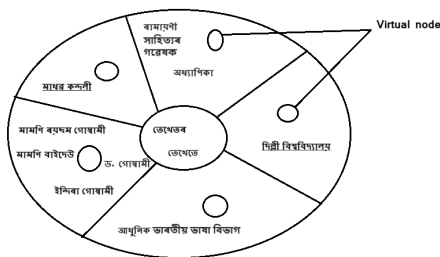


Fig. 5. Virtual Node Concept

Following Das et al. (2025) [5], we treat each mention (nouns and pronouns) as a node with associated attributes (syntactic, morphological, and lexical). Mentions that share semantic categories are grouped into partitions; a virtual node can be formed by merging attribute sets of two or more mentions. New mentions (including pronouns) query the virtual node when they attempt to link to an antecedent; acceptance is decided using rule checks.

**Original rule set:** Das et al. (2025) proposed six rules for peer-to-peer based coreference resolution:

**Morphosyntactic Agreement** (person, number, gender, etc.)

**Semantic Compatibility** (antecedent must belong to correct semantic class)

**Recency Preference** (prefer most recent compatible antecedent)

**Syntactic Parallelism** (antecedent in similar grammatical role preferred)

**Discourse Coherence** (antecedent should maintain overall coherence)

**Saliency Weighting** (apply preference hierarchy: subjects > objects > others)

**Adapted rule set:** For this study, we implemented only three rules due to resource constraints:

**Rule 1- (Morphosyntactic Agreement).** The candidate antecedent must agree with the anaphor on features explicitly realized in Assamese pronouns, primarily number (singular/plural) and person (1/2/3), and, where applicable, honorific/animacy cues. If the agreement fails, the candidate is rejected.

**Rule 2- (Semantic Category Compatibility).** The semantic class of the candidate antecedent (person, place, organization, abstract/event or object) must be compatible with the pronoun use. For example, a pronoun used for a human (তেওঁ) prefers antecedents labeled "Person", while এইটো or ইয়াত (when used for an inanimate referent) would more likely point to "Object/Event".

**Rule 3- (Recency/Saliency Preference).** When multiple candidates pass the above checks, prefer the most recent compatible antecedent weighted by grammatical saliency (subjects preferred over objects), sentence distance (same sentence > previous sentence), and syntactic prominence.

The remaining rules (syntactic parallelism, discourse coherence, and saliency weighting) were not implemented because they require reliable parsers and richly annotated discourse corpora in Assamese, which are not currently available. Thus, we chose the three most linguistically robust and practically feasible rules for Assamese.

A step-by-step illustration is provided below, along with a concrete example in Assamese:

**Original text (Assamese) and English translation:**

ৰাহুল এজন শিক্ষক। তেওঁ আজি যোৱা বিষয়টোৰ ওপৰত এটা বক্তৃতা দিছে।

Rahul is a teacher. He gave a talk today on the passed topic.

**Resolution steps:**

Identify all possible mentions along with their

grammatical features: বাহুল (M1, Person, singular), তেওঁ (A1, pronoun, 3rd person singular, honorific possible).

Candidate antecedents for তেওঁ: M1=‘বাহুল’ (and any other mentions in preceding sentences if present).

Apply Rule 1 (agreement): M1 is Person, singular — matches তেওঁ -> pass.

Apply Rule 2 (semantic): M1 semantic category Person -> compatible -> pass.

Apply Rule 3 (recency): M1 is in the immediately preceding sentence -> selected as antecedent.

**Notes on limitations:** This simple rule set handles many prototypical cases but fails when (i) the antecedent is not overt (pro-drop or implicit), (ii) the antecedent is a complex noun phrase split across sentences, (iii) "it"-like references point to clauses/events, or (iv) agreement markers are ambiguous in the surface form.

### 3.2.2 Text-Based Question Prompt Technique

Recent advances in LLMs have shown that prompt engineering can be an effective way to guide model responses toward specific tasks. However, simple strategies such as zero-shot or few-shot prompting often fail to capture long-range dependencies required for accurate anaphora resolution, particularly in morphologically complex languages like Assamese. To overcome this limitation, we employed a text-based question prompt technique, where the model is explicitly asked about the antecedent of a pronoun within a given discourse.

In this technique, a discourse segment is presented to the model along with a targeted question that directs attention to the anaphora resolution task. The question is designed such that the answer must be derived from the context provided. Unlike implicit prompting methods, this explicit formulation enables the model to focus on linking pronouns with their antecedents.

In this study, we used the **text-based question prompt technique**<sup>9</sup>. An example of such a sample prompt is illustrated in Fig. 6. In this prompt technique, a text is provided along with a question

designed so that the answer can be derived either directly or indirectly from the given text.

<p><b>Article:</b> Mehul Choksi, a key accused in the Rs 13,500 crore Punjab National Bank loan fraud case, was arrested in Belgium on Saturday. New Delhi has pressed for his extradition from Belgium, so that the 65-year-old fugitive diamantaire can be tried in India.</p>
<p><b>Question:</b> What is the referent of "his"?</p>
<p><b>Response:</b> The referent of "his" in the sentence is <b>Mehul Choksi</b>, the 65-year-old fugitive diamantaire mentioned earlier in the text.</p>

Fig. 6. A sample text-based question prompt

<p><b>অনুবন্দ:</b> মামণি বসুম গোস্বামী (ইংৰাজী: Mamoni Raison Goswami, জন্ম:১৯৪২ - মৃত্যু:২০১১) হৈছে অসমৰ এগৰাকী জ্ঞানপীঠ বঁটা বিজয়ী সাহিত্যিক। তেখেতৰ প্ৰকৃত নাম আছিল ইন্দিৰা গোস্বামী। সত্তে মামণি বাইদেউ হিচাবে তেখেতক জনা যায়। তেখেতে মাতৃভাষা অসমীয়াৰ উপৰিও ইংৰাজী ভাষাতো সাহিত্যকৃতি সৃষ্টি কৰিছিল। বামণী সাহিত্যৰ গবেষক হিচাপে ব্ৰাহ্মণ ক্ৰন্দলীৰ ৰামায়ণৰ গৱেষণাত ডক্টৰেট ডিগ্ৰী লাভ কৰা ড. গোস্বামীয়ে দিল্লী বিশ্ববিদ্যালয়ৰ আধুনিক ভাৰতীয় ভাষা বিভাগৰ অধ্যাপিকা ৰূপে অৱসৰ গ্ৰহণ কৰিছিল।</p>
<p><b>প্ৰশ্ন:</b> What is the referent of "তেখেত"?</p>
<p><b>Response:</b> In the given Assamese paragraph, the pronoun "তেখেত" (pronounced tekhet) is a respectful third-person singular pronoun. Throughout the passage, "তেখেত" refers to: মামণি বসুম গোস্বামী (Mamoni Raison Goswami), whose real name is ইন্দিৰা গোস্বামী (Indira Goswami). So, the referent of "তেখেত" is Indira Goswami, also known as Mamoni Raison Goswami.</p>

Fig. 7. Anaphora detection on minimum (minimum text that contains antecedent) text

During our experiment, we encountered several issues with this prompting technique. The prompt consists of two components: a given text and a question. The answer to the question, or the context for it, is expected to be derived from the provided text. In our experiment, we found that providing a complete discourse as the input text yields optimal results (Fig. 8). Here, the complete discourse may refer to either the full dataset or a partial excerpt from it. Our dataset consists of stories and text provider for the prompt that contain the answer to the question, either directly or indirectly.

Example. Consider the Assamese text:

<sup>9</sup><https://www.promptingguide.ai/introduction/examples> (accessed on 14-04-2025)

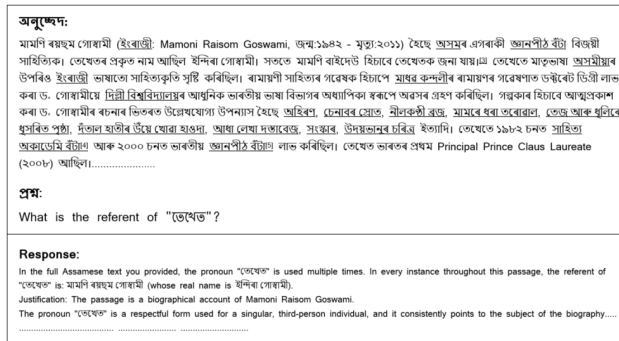


Fig. 8. Input the whole story to check for Anaphora in the entire text

“মামণি ৰয়ছম গোস্বামী হৈছে অসমৰ এগৰাকী জ্ঞানপীঠ বঁটা বিজয়ী সাহিত্যিক। তেখেতৰ প্ৰকৃত নাম আছিল ইন্দিৰা গোস্বামী।”  
 (“Mamoni Raisom Goswami is a Jnanpith Award winner literary figure from Assam. Her real name was Indira Goswami”)

Here, the pronoun “তেখেতৰ” (her) is the anaphor. The prompt provided to the model is: “In the above text, to whom does “তেখেতৰ” (her) refer?”

The expected answer is “মামণি ৰয়ছম গোস্বামী” (Mamoni Raisom Goswami). By framing the problem as a direct question, the model is guided to identify the correct antecedent instead of generating unrelated content.

This method offers several advantages in the Assamese context:

- It allows the model to capture long-range discourse dependencies, which are often missed by simple zero-shot or few-shot prompts.
- It provides robustness against morphological variations in Assamese pronouns, which can otherwise confuse rule-based approaches.
- It aligns closely with the cognitive process of human readers, who resolve anaphora by answering implicit “who/what” questions during comprehension.

Overall, the text-based question prompt technique proved more effective than the baseline rule-based system, demonstrating that carefully designed prompts can significantly improve anaphora resolution in Assamese.

### 4 Result and Analysis

In our approach, first, we use a minimum amount of text as input to the prompt for the anaphora resolution in the Assamese Language (Fig. 7). By minimal, we mean that the text must contain both the anaphora and its antecedent within a concise context. We use a text-based question prompt to resolve the anaphora. After that, we tried to input the whole discourse to check the anaphora (Fig. 8). The MUC score, a popular metric for assessing anaphora/coreference resolution, is used to present the results.

Table 2 shows a comparison of the results using a rule-based approach, i.e., peer to peer network for anaphora resolution and text based question prompt techniques. The MUC score of the rule-based system and prompt system is calculated. The MUC score reveals that the prompt-based approach has a higher F1-score as compared to the rule-based approach. When compared to the earlier attempts, the results shown in the comparison table (Table 2) are reasonably good.

Assamese coreference resolution research has recently achieved a MUC F1-Score of 0.7126 [5]. They used a tagged dataset with a rule-based approach and peer-to-peer network for coreference resolution to achieve the results. As anaphora is a subset of coreference, we use this rule-based approach for the resolution of pronominal anaphora. There isn’t any other study in prompt engineering that addresses pronominal anaphora in the Assamese language.

Therefore, we are unable to directly compare with other systems. To get better outcomes, we can say that the dataset and prompting strategies need to be improved. The algorithm will achieve improved outcomes with the use of a tagged dataset specifically designed for Assamese language questions. An error analysis is carried out in an effort to pinpoint the weaknesses in our system.

The following conclusion was reached in the case of Assamese language anaphora resolution, despite the analysis’s lack of thorough investigation: (i) More work needs to be done on prompting techniques; and (ii) In order to



**Table 2.** Comparison of results

MUC Score	Rule-Based System	Prompt System
Recall	0.617	0.729
Precision	0.746	0.856
F1-Score	0.674	0.787

comprehend the methods of pronominal anaphora resolution, the prompts should be trained using a large Assamese language test. Although our dataset is smaller, it has the most pronouns and was hand-picked by linguistic specialists. The prompt performs better with models that have a lot of data, but in our instance, using the Assamese language did not result in any erroneous data output. When we ask direct questions, the accuracy is higher. Research is still ongoing, and we are expanding our dataset.

## 5 Conclusion and Future Work

This paper presents a pioneering attempt to apply prompt engineering techniques for pronominal anaphora resolution in the Assamese language. While previous research in Indian languages has largely focused on rule-based or statistical approaches, our study introduces a novel text-based question prompt method with large language models and demonstrates its effectiveness in comparison with a peer-to-peer rule-based baseline.

The results indicate that the prompt-based technique achieves higher recall, precision, and F1-score than the rule-driven system, thereby establishing a strong benchmark for future work on Assamese discourse processing. The novelty of this study lies in three aspects: (i) the creation of a domain-diverse Assamese corpus rich in pronominal anaphora, (ii) the adaptation of a peer-to-peer rule-based method as a baseline for Assamese, and (iii) the introduction of a prompt-engineered approach specifically tailored to capture anaphoric relations in a morphologically complex, low-resource language.

Although the dataset used is relatively small, the findings provide compelling evidence that prompt engineering offers a viable and superior

alternative to traditional approaches for anaphora resolution in Assamese. Future research will focus on expanding the dataset, refining the prompting strategies, and developing annotated resources to enable more robust evaluation. We also envision that the methodology explored here can be extended to other low-resource languages, thereby contributing to the broader goal of inclusive natural language processing research.

## References

1. **Cherry, C., Bergsma, S. (2005).** An Expectation Maximization approach to pronoun resolution. Proceedings of the Ninth Conference on Computational Natural Language Learning (CoNLL-2005), pp. 88–95.
2. **Clark, K., Manning, C. D. (2015).** Entity-centric coreference resolution with model stacking. **Zong, C., Strube, M.**, editors, Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pp. 1405–1415. DOI: 10.3115/v1/P15-1136.
3. **Clark, K., Manning, C. D. (2016).** Deep reinforcement learning for mention-ranking coreference models. **Su, J., Duh, K., Carreras, X.**, editors, Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp. 2256–2262. DOI: 10.18653/v1/D16-1245.
4. **Das, M., Senapati, A. (2024).** Co-reference resolution in prompt engineering. *Procedia Computer Science*, Vol. 244, pp. 194–201. DOI: <https://doi.org/10.1016/j.procs.2024.10.192>. 6th International Conference on AI in Computational Linguistics.
5. **Das, M., Senapati, A. (2025).** Peer to Peer Network for Co-reference Resolution. *Journal of Scientometric Research*, Vol. 14, No. 1, pp. 342–350. DOI: 10.5530/jscires.20251457.

6. **Ge, N., Hale, J., Charniak, E. (1998).** A statistical approach to anaphora resolution. In Proceedings of the Sixth Workshop on Very Large Corpora, pp. 161–171.
7. **Hobbs, J. R. (1976).** Anaphora Resolution. Research Report 76-1. New York: Department of Computer Science, City University of New York.
8. **Itegulov, D., Lebedeva, E. (2018).** Handling verb phrase anaphora with dependent types and events. International Workshop on Logic, Language, Information, and Computation, Springer, pp. 210–222.
9. **Lalitha Devi, S., Sundar Ram, V., RK Rao, P. (2014).** A generic anaphora resolution engine for Indian languages. **Tsujii, J., Hajic, J.**, editors, Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, pp. 1824–1833.
10. **Mitkov, R. (1999).** Anaphora resolution: the state of the art. School of Languages and European Studies, University of Wolverhampton, pp. 1–34.
11. **Mitkov, R. (2014).** Anaphora Resolution. Routledge.
12. **Pakray, P., Neogi, S., Bhaskar, P., Poria, S., Bandyopadhyay, S., Gelbukh, A. (2011).** A textual entailment system using anaphora resolution. In: TAC 2011.
13. **Peng, H., Khashabi, D., Roth, D. (2015).** Solving hard coreference problems. In Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 809–819.
14. **Peng, H., Khashabi, D., Roth, D. (2019).** Solving hard coreference problems. arXiv preprint arXiv:1907.05524.
15. **Plu, J., Prokofyev, R., Tonon, A., Cudré-Mauroux, P., Difallah, D. E., Troncy, R., Rizzo, G. (2018).** Sanaphor++: Combining deep neural networks with semantics for coreference resolution. LREC 2018, 11th Language Resources and Evaluation Conference.
16. **Poesio, M., Kabadjov, M. A. (2004).** A generalpurpose, off-the-shelf anaphora resolution module: Implementation and preliminary evaluation. International Conference on Language Resources and Evaluation.
17. **Pradhan, S., Moschitti, A., Xue, N., Uryupina, O., Zhang, Y. (2012).** CoNLL-2012 shared task: Modeling multilingual unrestricted coreference in OntoNotes. Joint Conference on EMNLP and CoNLL - Shared Task, pp. 1–40.
18. **Rahman, A., Ng, V. (2012).** Resolving complex cases of definite pronouns: The Winograd schema challenge. Proceedings of the 2012 Joint Conference on EMNLP and Computational Natural Language Learning, pp. 777–789.
19. **Rich, E., LuperFoy, S. (1988).** An architecture for anaphora resolution. Proceedings of the Second Conference on Applied Natural Language Processing (ANLP-2), pp. 18–24. DOI: "10.18653/v1/2021.naacl-main.329.
20. **Senapati, A., Garain, U. (2013).** GuiTAR-based pronominal anaphora resolution in Bengali. **Schuetze, H., Fung, P., Poesio, M.**, editors, Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pp. 126–130.
21. **Sikdar, U. K., Ekbal, A., Saha, S., Uryupina, O., Poesio, M. (2015).** Differential evolution-based feature selection technique for anaphora resolution. *Soft Computing*, Vol. 19, pp. 2149–2161. DOI: <https://doi.org/10.1007/s00500-014-1397-3>.
22. **Sobha, L., Patnaik, B. N. (2000).** VASISTH: an ellipsis resolution algorithm for Indian languages. Proceedings of the International Conference on Machine Translation and Multilingual Applications in the new Millennium: MT 2000.

23. **Soon, M. W., Lim, C. Y. D., Ng, T. H. (2001).** A machine learning approach to coreference resolution of noun phrases. *Computational Linguistics*, Vol. 27, No. 4, pp. 521–544.
24. **Versley, Y., Ponzetto, S., Poesio, M., Eidelman, V., Jern, A., Smith, J., Yang, X., Moschitti, A. (2008).** BART: A modular toolkit for coreference resolution. *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, pp. 9–12.
25. **Wiseman, S., Rush, A. M., Shieber, S. M. (2016).** Learning global features for coreference resolution. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 994–1004. DOI: 10.18653/v1/N16-1114.
26. **Yang, X., Peynetti, E., Meerman, V., Tanner, C. (2022).** What GPT knows about who is who. **Tafreshi, S., Sedoc, J., Rogers, A., Drozd, A., Rumshisky, A., Akula, A.**, editors, *Proceedings of the Third Workshop on Insights from Negative Results in NLP*, pp. 75–81. DOI: 10.18653/v1/2022.insights-1.10.
27. **Yu, J., Moosavi, N. S., Paun, S., Poesio, M. (2021).** Stay together: A system for single and split-antecedent anaphora resolution. *Association for Computational Linguistics*, pp. 4174–4184. DOI: "10.18653/v1/2021.naacl-main.329.

*Article received on 30/05/2025; accepted on 05/09/2025.*  
*\*Corresponding author is Mridusmita Das.*