International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

A Novel Anaphora Resolution Model for Afaan Oromo Language Texts

Misganu Tuse Abetu [1], Ramesh Babu P [2]
[1 2] College of Engineering & Technology, Wollega University,
Post Box No: 395, Ethiopia

***Abstract*:- In Natural Language, Anaphora is an expression which refers to another expression in its context. This referred expression, called the antecedent provides the information for interpretation of the anaphora. Anaphora Resolution is the task of identifying the antecedent of the anaphora. Anaphora resolution is required in various NLP applications such as information extraction, summarization and Machine translation. Anaphora resolution is used as a component in NLP applications like machine translation, information extraction, question answering and others to increase their effectiveness.**

The native name for the Oromo language is Afaan Oromo, and it is an official language in Ethiopia. This proposed model takes Afaan Oromo texts as input and preprocesses to tag the texts with word classes and various chunks. Anaphors, both independent and hidden, and antecedents are identified from the preprocessed dataset. The model deals with both intra-sentential and inter-sentential type of anaphors. In this paper, we have proposed Afaan Oromo anaphora resolution model using knowledge poor anaphora resolution approach. Afaan Oromo texts are collected from BBC Afaan Oromo and Afaan Oromo Holy Bible to evaluate the performance of this model. The collected dataset was divided into training and testing datasets based on 10-fold cross validation technique. Based on the collected dataset, we achieved precision of 80.53%, recall of 90.96% and F-measure of 85.22% for resolution of independent anaphors whereas precision of 87.3%, recall of 92.16% and F-measure of 89.56% was obtained for resolution of hidden anaphors.

Keywords: Afaan Oromo, Anaphora resolution model, knowledge poor approach, hidden anaphors and texts.

1. INTRODUCTION

NLP is an area of computer science and artificial intelligence concerned with the interactions between computers and human (natural) languages. in particular how to program computers to fruitfully process large amounts of natural language data Natural Language Processing aims at designing and building software that will analyze, understand, and generate human languages, so that eventually humans will be able to communicate with the machines using natural languages. The word anaphora came from two ancient Greek words “ana” and “phora”. “Ana” means back, upstream, back in an upward direction whereas “phora” means the act of carrying (Ruslan miktov, 1999). So Anaphora means a word or phrase that refers to an entity that is mentioned previously and this word or phrase that it refers to is called its antecedent. It is a technique or phenomena of pointing back to an entity that has been introduced with more descriptive phrase in the text than the entity or expression which is referring back. The entity referred in the text can be anything like object, concept, individual, process, or any other thing. An entity referring back is called anaphor, whereas the previous expression being referred is called antecedent (Pilleriin Mutso, 2008).

The process of finding the antecedent for an anaphor is Anaphora Resolution. The relation exists between an anaphor and antecedent increases cohesiveness of sentences because it is used frequently in both written and oral communications to avoid over repetition of terms. The correct interpretation of anaphora is vital for Natural Language Processing.

See the following examples to get clear understanding of anaphora resolution.

1. <u>Tolaan</u> gara Gimbii <u>deeme.</u>

In this example the verb “deeme” is anaphour referring to the antecedent “Tolaa”.

2. <u>Caalaan</u> baayee hamaadha. <u>Inni</u> yeroo hundaa hammina barbaada.

In this example the pronoun “inni” is anaphour referring back to the antecedent “Caalaa”.

3. Yoseef <u>obboloota</u> isaa barbaacha deeme. <u>Obbolootni</u> isaas inni gara isaanii osoo dhufaa jiruu fagootti isa <u>argan.</u>

In this example the verb “argan” is anaphor referring back to antecedents “obboloota Yoseef”.

4. <u>Dubartittin</u> firiin mukichaa nyaataaf bareedaa akka ta’e gaafa agarte, kutattee nyaatte. <u>Isheen</u> isa ishee faana isa jiruufis akka nyaatuuf laatteef, Innis Ni nyaate.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

In the above example the pronoun "Isheen" is anaphour referring back to the antecedent "dubartittii" in the above sentence. To identify these kinds of hidden anaphors it needs morphological analyzer which makes Afaan Oromo anaphora resolution process more complex than other languages.

This paper is organized as follows, Section 1 contains the introduction of Cyber forensics and image retrieval system, Section 2 contains the literature survey of various image retrieval methods, Section 3 contains the brief description of materials, methods and some test results Section 4 describes about results and discussions Section 5 concludes about this proposed Anaphora model.

2. LITERATURE REVIEW

2.1. Anaphora Resolution Approaches

The formalization of world knowledge, various semantic issues and other issues which are important for the implementation of successful anaphora resolution systems are still far from being achieved.

As a result, in 1990s most researches were shifted to exploiting lower-level information like morphology and syntax from the higher level information which are complex. So, based on the level of information required to resolve anaphora, two broad categories have emerged: knowledge-rich and knowledge-poor approaches. Both approaches are discussed in the following sub sections (Dilek Küçük, 2005).

2.1.1. Knowledge-Rich Approaches

Knowledge rich anaphora resolution approaches are approaches that uses knowledge such as morphological, syntactic, semantic, discourse, domain knowledge which are needed to deal with anaphora resolution using knowledge rich approaches (Tejaswini Deoskar,2001). Totally this approaches are approaches that employ linguistic and domain knowledge in great detail for anaphora resolution process starting from the morphological knowledge to the very high level knowledge like world knowledge.

These approaches are rule based and it can be further divided in to four: factor based approaches, discourse oreinted approaches, heuristic based approaches and syntax based approaches (Caroline V. Gaspering, 2008).

2.1.2. Knowledge-Poor Approaches

Knowledge-poor anaphora resolution algorithms are the algorithms which do not use too much linguistic and domain knowledge to resolve the anaphors. This approaches are formulated to avoid the complex syntactic, semantic, discourse and world knowledge‟s used in anaphora resolution process. Since knowledge based approaches are labor intensive and time consuming plus computationally expensive, knowledge poor anaphora resolutions are formulated to solve these complexities. It is a result of high need for inexpensive solutions to satisfy the need of NLP systems in practical way. The development of NLP tools like POS tagger and other tools facilitated for the invention of knowledge poor approach. This approach makes use of part of speech tagger, noun phrase rules and then applies antecedent indicators which are constraint and preference rules on a set of potential antecedents. Knowledge poor anaphora resolution approach depends on a set of constraint and preference rules (Ruslan Miktov, 1999).

Anaphora resolution approaches that showed the practicality of knowledge poor anaphora resolution approach include: Kennedy and Boguraev‟s approach without a parser (Ruslan Miktov, 1999). Robust knowledge poor approach (R.Mitkov, 1998) (Ruslan Mitkov, 1999).Collocation patterns-based approach (Tejaswini Deoskar & Ruslan Miktov, 2001).

2.2. Types of anaphora

In addition to the knowledge needed to perform anaphora resolution, the various forms that anaphora can assume make it a very challenging task to teach computers how to solve anaphora. Because different anaphora type may be different in structure because of its position in the sentence. When we see the forms of anaphora, Anaphora can be divided into pronominal anaphora, lexical noun phrase anaphora, noun anaphora, verb anaphora, adverb anaphora or zero anaphora based on the form of the anaphor or syntactic category of the anaphor. It can also be classified into intrasentential or intersentential based on the location of the anaphor (Vaclav Nemcik, 2006) (Dilek Küçük, 2005) (Temesgen D, 2014).

3. MATERIALS AND METHODS

Methodology is a way to systematically solve the research problem (Lappin et al., 1994). The methodology focuses on the approaches, techniques, tools and data source the researcher followed during the course of the study. This section is very imperative as the performance and efficiency of the final system depends on the validity and reasonable approaches and techniques followed by the researcher. Accordingly, the tools, study design and others that the researcher used are detailed below.

3.1. Development Tools and Techniques

In the development of Afaan Oromo Anaphora resolution model the tools we have used is **Anaconda Programming Language** the reason why Anaconda is selected is that it has rich implementations of the widely used NLP algorithms and Libraries including python. It is a powerful language and has the right combination of performance and features that make writing programs in Anaconda both fun and easy. It is dynamic programming language that is used in a wide variety of application domains.

We used Knowledge-poor anaphora resolution approach since it do not use too much linguistic and domain knowledge to resolve the anaphors. This approaches are formulated to avoid the complex syntactic, semantic, discourse and world knowledge used in anaphora resolution process.

3.2. Study Design

This research is an experimental design that tests the implementation of anaphora resolution model for Afaan Oromo texts. The proposed model has two parts the first is resolution of hidden anaphors and the second one is resolution of independent anaphors.

3.2.1. Architecture of Afaan Oromo Anaphora Resolution Model

Proposed Afaan Oromo AR architecture has two parts: resolution of independent anaphors and resolution of hidden anaphors. It shows components used to resolve independent Afaan Oromo personal pronouns (independent anaphors) and hidden Afaan Oromo personal pronouns (hidden anaphors).

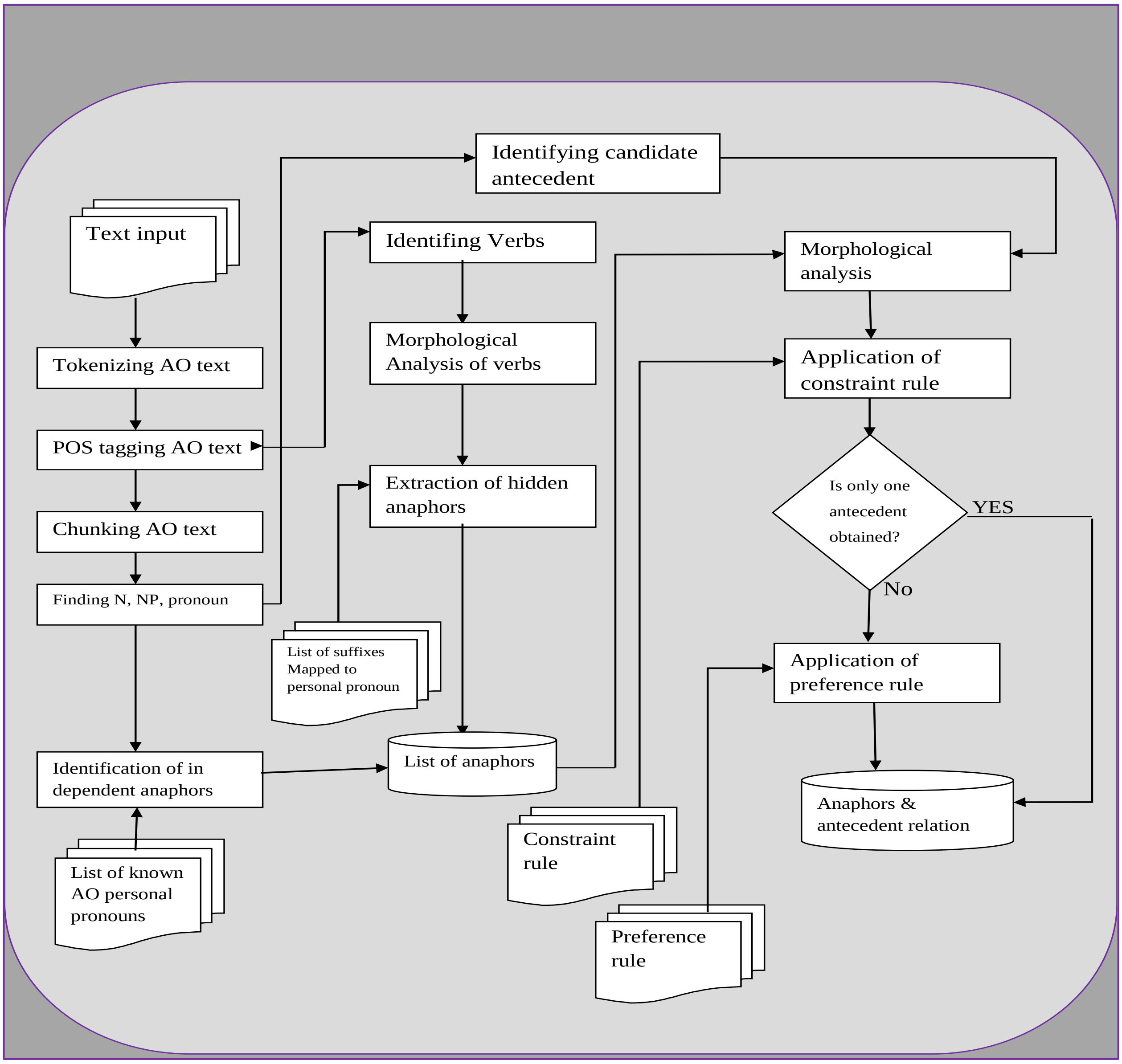

Figure 3.1.Architecture of Anaphora resolution model for Afaan Oromo texts

3.2.2. 3.2.1 Preprocessing steps

Tokenization

Tokenization is the process of dividing the given text or sentence in to tokens and tokens are may be word, phrase, characters or other important unit by using predefined list of delimiters such as new lines and space. The example of tokenization is as follows

Input: Tolaan gara Gimbii deeme.

Output: [Tolaan | gara | Gimbii | deeme]

Afaan Oromo POS Tagging.

We used POS tagging because it helps identify antecedents and independent anaphors, and used as input for chunking. Because of the unavailability of working POS tagger for Afaan Oromo, we tagged Afaan Oromo sentences manually.

Chunking

Chunking is helpful for those antecedent that comprise two or more words. For example, let us consider the name "mana barumsaa" which is composed of two words.if we take the output of POS tagger "mana " is tagged as NN and "Barumsaa" is also tagged as NN but both words follow each other and also belong to the same thing. We chunked Afaan Oromo sentences manually for the resolution of anaphors since there is no chunker for Afaan Oromo.

Example

- ✓ Bulchiinsi Magaalaa Naqamtee Gamoo Bareedaa ijaare

[('Bulchiinsi N magaalaa N naqamtee N','NP'), ('Gamoo N Bareedaa N ijaare V', VP)].

Finding N, NP and Pronoun

From the chunked Afaan Oromo sentences both Noun and Noun phrase are taken as candidate antecedent for the resolution of hidden anaphors. And getting pronoun is used for the identification of independent anaphors.

Identification of independent anaphors

Independent Afaan oromo personal pronouns are considered as independent anaphors in anaphora resolution.

Identification of hidden Anaphors

Hidden anaphors means personal pronouns found inside a given verb. Sometimes Afaan Oromo verbs are considered as a sentence.

Example

Deemaniiru: the hidden anaphors inside the verb "deemaniiru" is a personal pronoun "Isaan".

Identification of verbs

Identification of verbs is needed to extract personal pronouns hidden inside them. As a result, the POS Tagging result previously generated in the Data Processing component is used to identify verbs.

Morphological analysis of verbs

After verbs are identified, their morphology is analyzed so that hidden anaphors are extracted.

The following examples show morpheme formation (person, number, gender) of verbs.

- ✓ Bite: bit-e = 3rd person ,singular,male
- ✓ Bitte:bit-t-e = 3rd person ,singular, Female
- ✓ Bitan: bit-an= 3rd person, plural.
- ✓ Bitne:bit-n-e =1st person ,plural
- ✓ Bitti:bit-t-i = 3rd person ,singular, Female
- ✓ Bita:bit-a = 1st person,3rd person ,singular,male
- ✓ Bitu:bit-u = 3rd person ,plural
- ✓ Bitna:bit-n-a = 1st person ,plural

Extraction of hidden Anaphors

Extraction of hidden anaphor is the process of finding the corresponding Afaan Oromo personal pronoun for a given verb.

Selection of candidate antecedent

All nouns and head nouns of noun phrases are used as possible candidate antecedents. Candidate antecedents is list of antecedent compete to be the correct antecedents referred by anaphors are selected from it. So, this component is used to find possible antecedents which could be best antecedent of the anaphor to be resolved.

Resolution process

Anaphora resolution is about finding correct antecedent for the given anaphors. So, this component resolves anaphors to their correct antecedents based on identified candidate antecedents and anaphors. The subcomponents of this component are morphological analysis, application of constraint rules and application of preference rules.

Morphological Analysis

Morphological analysis is one step of anaphora resolution since the information such as Gender, number and person information of words are bound morphemes in Afaan Oromo. To identify those information, morphological analysis is needed.

Application of constraint rules

The constraint rules used in this work are gender, number and person agreements. They are described as follows:

i. Both anaphors and antecedents should agree in gender information
ii. Both anaphors and antecedents should agree in number information
iii. Both anaphors and antecedents should agree in person information.

Application of preference rule

After removing incompatible candidates, if the remaining list contains more than one antecedent, preferences are applied in order to choose a single antecedent.

Preference rules used in this works are Subject place, definiteness, boost pronoun and highest frequency.

3.3. Development of Dataset

Sentence having independent personal pronoun was collected from Afaan Oromo holly bible and sentences we have used for resolving hidden Afaan Oromo was collected from BBC Afaan Oromo hence the dataset we used in this work was collected from two sources. The first set of dataset used for extraction of hidden anaphor. Sentences having independent personal pronouns are collected from Afaan Oromo holly bible. The collected sentences from the Afaan Oromo holly bible are manually tagged and chunked. A total of Design of Afaan Oromo Anaphora Resolution that having **150** personal pronouns are collected from Afaan Oromo holly bible. A total of **250** sentences, collected in order to resolve hidden anaphors, are used to decide the values of preference rules. Out of the total sentences collected, 135 sentences having antecedents at subject place, 80 sentences having anaphors referring back to antecedents which are definite, 55 sentence having antecedent which are personal pronouns and 10 sentences where nouns mentioned more than once are selected as antecedents are identified. Each of the preference rule values are calculated manually using probability of distribution they have based on the collected sentences. From the collected sentences, most of the personal pronouns refer to nouns at subject places, the second is definite noun, Boost pronouns are the third largest set of class referred by personal pronouns and lastly nouns which are frequent are ranked third to be referred by personal pronouns.

3.4 *Test results*

A technique called 10-fold cross validation is used to prepare the dataset as training and testing. In 10-fold cross validation, the entire dataset is randomly partitioned into k equal samples. From these k samples, one sample is used as testing dataset whereas the remaining k-1 samples are used as training datasets. Since it is 10-fold cross validation, the process is repeated 10 times. After preparation of the dataset for testing and training, the next step we performed was evaluating performance of the model.

The formula for calculating precision, recall and F-Score are as mentioned below:

- $$\text{Precision} = \frac{\text{Correctly resolved}}{\text{Correctly resolved} + \text{wrongly resolved}} \quad \ldots\ldots 4.1$$

- $$\text{Recall} = \frac{\text{Correctly resolved}}{\text{Correctly resolved} + \text{wrongly rejected}} \quad \ldots\ldots 4.2$$

- $$\text{F-Measure} = \frac{2*(\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}} \quad \text{……………………..………………..4.3}$$

Performance of AOAR model for independent anaphors

We have collected 150 sentences for resolving independent anaphors. From this sentences there is anaphor which is correctly resolved (true positive), wrongly resolved (False positive) and wrongly rejected (False negative) as we have seen from the following table. For independent anaphors has the overall performance of 80.53% precision, 90.96% recall and 85.22% F-measure in terms of percentage.

Table 4.1 result of AOARM on independent Anaphors

Dataset	**Correctly resolved**	**Wrongly resolved**	**Wrongly rejected**	**Correctly rejected**	**Precision**	**Recall**	**F-measure**
Dataset 1	9	4	2	0	0.692	0.810	0.746
Dataset 2	13	2	0	0	0.867	1	0.929
Dataset 3	12	3	0	0	0.8	1	0.889
Dataset 4	11	3	1	0	0.786	0.917	0.846
Dataset 5	10	3	2	0	0.769	0.833	0.799
Dataset 6	12	2	1	0	0.857	0.923	0.889
Dataset 7	11	4	0	0	0.733	1	0.846
Dataset 8	12	1	2	0	0.923	0.857	0.889
Dataset 9	10	3	2	0	0.769	0.833	0.799
Dataset10	12	2	1	0	0.857	0.923	0.889

Performance of AOAR model for anaphors hidden inside verbs

Our AOARM for hidden anaphors has the overall performance of 87.30% precision, 92.16% recall and 89.56% F-measure in terms of percentage.

Table 4.2 result of AOARM on hidden Anaphors

Dataset	**Correctly resolved**	**Wrongly resolved**	**Wrongly rejected**	**Correctly rejected**	**Precision**	**Recall**	**F-measure**
Dataset 1	21	2	2	0	0.913	0.913	0.913
Dataset 2	18	4	3	0	0.818	0.857	0.837
Dataset 3	22	2	1	0	0.917	0.957	0.936
Dataset 4	19	5	1	0	0.792	0.95	0.864
Dataset 5	20	2	2	1	0.909	0.909	0.909
Dataset 6	17	4	3	1	0.810	0.85	0.829
Dataset 7	23	2	0	0	0.92	1	0.958
Dataset 8	20	3	2	0	0.870	0.909	0.889
Dataset 9	19	4	1	1	0.826	0.95	0.884
Dataset10	22	1	2	0	0.957	0.917	0.936

4. RESULT AND DISCUSSIONS

In the discussion we deal with some events which leads to reduction in performance of Afaan Oromo Anaphora resolution model.

The following points illustrate the events or the condition in which AOARM cannot resolve correctly:

1. There is a condition were verb and noun mismatch in gender agreement with in a sentence. Consider the following example

Example

Itoopiyaan hariiroo hawaas-dinagdee Eritriyaa waliin qabdu cimsite.

As we see from the above example the verb ‘cimsite’ has personal pronoun ‘Ishee’ (she) hidden inside it. ‘Itoopiyaa’ is correct antecedent from existing choices preceding the verb and should also be selected as best antecedent by Afaan Oromo AR model. But Afaan Oromo AR model resolution process didn’t select ‘Itoopiyaa’ as correct antecedent. It is because when the constraint rules such as gender, number and person information, are applied it didn’t match and as a result discarded. Gender, number and person information of ‘Itoopiyaa’ is 3rd person singular masculine. Gender, number and person information of the verb ‘cimsite’ is 3rd person singular and having gender information female.

2. When we going to resolve personal pronouns hidden inside verbs, we have found sentences talking about someone in singular way showing politeness. On these sentences verbs are represented as plural but semantically they are not plural rather they show

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

politeness. Whereas, nouns that they are referring to are presented as singular. As a result, they don't match number agreement when constraint rules are applied. See the following examples.

Example,

Obbo Lammaan docteera kabajaa yuunivarsiitii jimmaa irraa fudhatan.

From the above example consider the verb 'fudhatan' it is 3rd person plural but the antecedent of this sentence is 'obbo Lammaa' which is 3rd person singular so here there is mismatch in number agreement between the anaphor and the correct antecedent. Because of these our Afaan Oromo Anaphora resolution model cannot resolve such sentence since they didn't mismatch person agreement rather they show politeness.

3. There is a condition where there is no antecedent in a sentence that described by anaphor while the sentence valid.

Example

Caalaan maalif har'a akka mana barumsaa hafe beeku.

From the above example if we consider the verb 'beeku' it is 3rd person plural (they) but there is no antecedent referred by this anaphor while the sentence is valid.

5. CONCLUSION

In this paper the proposed Afaan Oromo Anaphora resolution model has two parts. The first part is resolving the hidden anaphor's inside verb and the second part is resolving independent anaphor i.e. Afaan Oromo personal pronouns. To extract hidden anaphors we take verb from Afaan Oromo sentence and then identifying morphological information of the verb.

The developed Afaan Oromo Anaphora resolution model has some component such as data preprocessing, identification of hidden anaphors, identification of independent anaphors, identification of candidate antecedents and Anaphora resolution. To evaluate the performance of the AOARM we have collected 250 sentences for resolving hidden anaphors inside verb and 150 for resolution of independent anaphors. As we have described in the test results we used parameters such as precision, Recall and F-measure. Accordingly AOARM for independent anaphor performed precision of 80.53%, Recall 90.96% and 85.22% F-measure. Whereas Performance of AOARM for hidden anaphors inside verb has precision 87.30%, Recall 92.16% and 89.56 F-measure in terms of percentage.

6. REFERENCES

[1] Hobbs (1977). Resolving pronoun references, in Readings in Natural Language Processing, Grasz, Jones and Webber, eds., Morgan Kaufman Publishers, Inc. Los Altos, California, USA.
[2] Bobrow, D.G (1964). A question-answering system for high school algebra word problems. Proceedings of AFIPS conference.
[3] Caroline V. Gaspering (2008). "Statistical Anaphora Resolution in biomedical texts".In Proceedings of the 22nd International Conference on Computational Linguistics, Stroudsburg, PA, USA: 257-264.
[4] Ruslan Miktov (1999). "Anaphora Resolution: The State of the Art," Working paper (Based on the COLING'98/ACL'98 tutorial on anaphora resolution).
[5] Vaclav Nemcik (2006). "Anaphora Resolution", Master‟s Thesis, Masaryk University, Brno.
[6] Tilahun Gamta (1992). Professor of Linguistic Studies at the University of Addis Ababa and author of Oromo/English dictionary
[7] Kohavi, R (1995). "A study of cross-validation and bootstrap for accuracy estimation and model selection" In Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence)
[8] Niklas Jakob and Iryana Gurevych (2010). "Using Anaphora Resolution to Improve Opinion Target Identification in Movie Reviews", In Proceedings of the ACL 2010 conference short papers, Sweden: 263-268.
[9] Shivangi shekhar and umesh kumar (2002). Review on the techniques of anaphora resolution
[10] Lappin, Shalom, Herbert Leass (1994). "An algorithm for pronominal anaphora resolution", Computational Linguistics: 535-561.
[11] Michael Gasser (2011). "HornMorpho: a system for morphological processing of Afaan Oromo, Oromo, and Tigrinya", Conference on Human Language Technology for Development, Alexandria, Egypt.
[12] Dilek Küçük (2005). "A knowledge-poor pronoun resolution system for Turkish", A thesis submitted to the Graduate School of Natural and Applied Sciences of Middle East Technical University.
[13] Kabadjov (2007). "A Comprehensive Evaluation of Anaphora Resolution and Discourse-new Classification", PhD thesis, Department of Computer Science, University of Essex.
[14] Pilleriin Mutso (2008). "Knowledge-Poor Anaphora Resolution System for Estonian", Master‟sThesis, University of Tartu, Tartu.
[15] Nuno Nobre (2011). "Anaphora resolution", Master‟s thesis, Instituto SuperiorTecnico / Universidade Tecnica de Lisboa.
[16] Temesgen D. (2014). "Knowledge poor Amharic anaphora Resolution model" Addis Abeba, Ethiopia.
[17] Tejaswini Deoskar (2001). Techniques for Anaphora Resolution: A Survey.
[18] Kennedy, Christopher, Branimir Boguraev (1996). "Anaphora for everyone: pronominal anaphora resolution without a parser". Proceedings of the 16th International Conference on Computational Linguistics (COLING'96), 113-118. Copenhagen, Denmark
[19] Ruslan Mitkov, Richard Evans, Constantin Orasan (2002). "A new, fully automatic version of Mitkov‟s knowledge-poor pronoun resolution method", In Proceedings of the third International conference on Intelligent Text Processing and Computational Linguistics , Mexico city, Mexico.

[20] Holen (2006). Automatic Anaphora Resolution for Norwegian (ARN) OSLO university department of of Linguistics and Scandinavian Studies
[21] R.Mitkov (1998). “Robust Pronoun Resolution with limited knowledge”, In Proceedings of the 36th Annual Meeting of the Association for computational Linguistics and 17th International Conference on Computational Linguistics, Montreal, Canada: 869-875.
[22] Baldwin (1997). CogNIAC: High precision coreference with limited knowledge and linguistic resources. In Proceedings of the ACU97/EACU97 Workshop on Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts
[23] Jurafisky and Martin (2000) Speech and Language Processing an Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition
[24] Priyanka Ahir (2006). Student of M.Sc. (IT), 2Faculty of M.Sc. (IT) 1,2Uka Tarsadia University Bardoli, Surat. Automatic Anaphora Resolution for Norwegian (ARN) hybrid approach
[25] Tyne Liang and Dian-Song Wu (2004). “Automatic Pronominal Anaphora Resolution in English Texts”, Computational Linguistics and Chinese Language Processing.
[26] Silvestre Marques (2013) anaphora Resolution in PortugueseAn hybrid approach department of information system and computer engineering.
[27] R.Mitkov (1998). “Robust Pronoun Resolution with limited knowledge”, In Proceedings of the 36th Annual Meeting of the Association forcomputational Linguistics and 17th International Conference on Computational Linguistics.