International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

Analyzing Coreference Tools for NLP Application

Sandhya Singh[1*], Krishnanjan Bhattacharjee[2], Hemant Darbari[3], Seema Verma[4]

[1,4]Banasthali University, Rajasthan, India
[2,3]C-DAC, Pune, Maharashtra, India

[*]*Corresponding Author: sandhya.singh@gmail.com*

DOI: https://doi.org/10.26438/ijcse/v7i5.608615 | Available online at: www.ijcseonline.org

Accepted: 13/May/2019, Published: 31/May/2019

Abstract— Coreference resolution is an important processing step for semantic analysis of a text in NLP. It facilitates in better understanding of the text. So coreference resolution tool becomes a necessity for every NLP process meant for text understanding or generation. The task of selecting a tool from a range of available open source coreference resolution tools can be challenging. This paper presents a study of these available open source coreference resolution tools with the aim to select a better performing tool that can be integrated into an NLP pipeline with ease. After the initial theoretical study of 13 open source coreference tools, a black box testing approach has been followed for testing the performance of 5 selected tools for their performance, usage and ease of integration for building an NLP application like summarization system, dialogue system etc. The performance evaluation is done using standard CoNLL 2012 coreference dataset for English language. The coreference marked output is evaluated against the manually tagged gold standard dataset. The performance is analyzed to select the best performing coreference tool for practical applications.

Keywords— Coreference resolution, coreference tools, entity resolution, NLP

I. INTRODUCTION

With a huge amount of textual information openly available on the web, the need for Automatic Text Summarization (ATS) and Information Extraction (IE) in particular are becoming more relevant. These applications of NLP require automated processing at morphological, syntactic, semantic and pragmatic levels of language. For understanding any sequence of sentences as a coherent text, or in discourse interpretation, coreference has a major role to play at the linguistic level. They represent the agreement between sentences and act as a link among sentences to show the cohesion and coherence thus making it easier to comprehend.

The literature clearly indicates that the thrust in the area of coreference resolution tools and its automatic evaluation techniques started in the early 1990s with MUC-6 conference [1] and reached a reasonably stable state in the terms of standard evaluation metrics and multiple open source quality tools for usage by 2012 with CoNLL-2012[2] conference shared task. Though, most of the tools available for usage were designed and developed during this interval, the problem of coreference resolution is still an open problem. The progress in the approach used for resolving coreference is evident from the range of features used in resolving coreference to the availability of knowledge resources from where these features were being extracted.

On the tool development front, the algorithm approach has progressed from supervised to unsupervised machine learning and is moving towards deep neural network approaches.

This paper presents a systematic study of 13 openly available coreference tools and libraries from different NLP labs. This study was extended to choose the better performing coreference tool for the purpose of integrating it to an NLP application. From the tools studied, 5 were further selected for black box testing to get a hands-on usage experience and performance analysis using CoNLL-2012 gold standard dataset for coreference resolution.

The rest of the paper is organized as follows: section II explains coreference resolution as a kind of entity resolution and its related terminology, section III discusses previous work and surveys about coreference and how this survey study is able to fill the gap. Section IV focuses on the theoretical study of a list of 13 coreference tools and libraries. Section V further analyses the performance of 5 tools with reference to CoNLL-2012 gold standard dataset for English language, section VI discusses the result and section VII concludes the paper.

II. UNDERSTANDING COREFERENCE

In linguistics, discourse knowledge is included in the text through entity resolution and linking. Entity resolution refers

[1] https://cs.nyu.edu/cs/faculty/grishman/muc6.html
[2] http://emnlp-conll2012.unige.ch/

to identifying named entities viz. name, place, organization and resolving it with pronouns/anaphora in place of repeated occurrences of the entity in a discourse. The entity resolution process involves recognizing, resolving and linking the mentions. Research has shown that NLP applications like machine translation, text summarization, information extraction, paraphrase detection etc. have gained improvement due to the entity resolution process. The research in entity resolution has been focussed on anaphora resolution and coreference resolution mainly.

Anaphora is the expression whose interpretation depends upon another expression in its antecedent. The process of finding this antecedent for an anaphor is known as anaphora resolution. Here, the anaphor is the reference pointing to the previous item while antecedent is the entity which the anaphor is referring to. *Example 1* below shows a sentence with *'It'* as anaphor referring to the antecedent *'the car'*.

Example 1: *"She saw [the car] going on the road. [It] later crashed into a roadside tree."*

Some common anaphor types are pronominal anaphora, definite noun phrase anaphora, verb anaphora *etc.* Pronominal anaphora is the most experimented anaphora in NLP.

Coreference is the mentions referring to the same real-world entity. Coreference resolution is the task of finding the mentions in text and clustering them according to the entity they refer to. Here, an entity denotes an object in the world and mention is a referring expression that describes that object. *Example 2* below shows a coreference resolution cluster containing *'Mohandas Karamchand Gandhi', 'He'* and *'Gandhi Ji'* where *'Mohandas Karamchand Gandhi'* is an entity and *'He' and 'Gandhi Ji'* are mentions.

Example 2: *"[Mohandas Karamchand Gandhi] was a great freedom fighter. [He] was regarded as the " Father of the Nation". [Gandhi Ji] believed in satya and ahimsa."*

In a text, a specific anaphor and some preceding or following noun phrases may be coreferential, thus forming a coreference chain of entities which have the same referent. Coreference is a specific case of anaphor realized by pronouns and non-pronominal definite noun phrases. *Figure 1* shows the relation between anaphora resolution and coreference resolution [1].

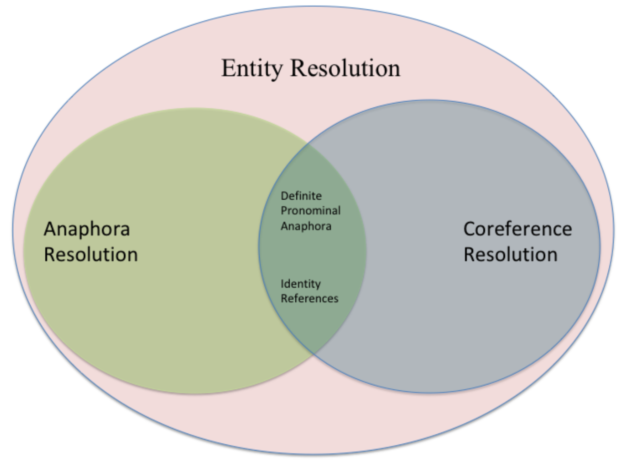

Figure 1: Relation between anaphora and coreference [1]

Coreference resolution is an important processing step in NLP for implementing coherence and cohesion as machines can process language better by substituting all the references with their entities. when multiple terms in a text point to the same entity in the real world, a chain is formed which helps in sequencing the text for better understanding. Linguistically, coreference can be of pronominal, nominal or a named entity type. *Figure 1* shows that definite pronominal anaphora and identity references are special kind of coreference represented using pronouns thus making it an anaphora as well. Coreference can exist at an inter-document level as well but not anaphora. So computationally, the task of coreference resolution searches for all forms of coreference categories like a noun phrase, pronouns *etc.*, whereas anaphora resolution, requires to focus on identifying all forms of the antecedent referring to the same sense.

At a human level, people are tuned to automatically resolve coreference by using the context and grammar clues along with world knowledge and meaning derived from the text. But implementing it computationally is a challenging task, as it requires a good amalgamation of semantics and world knowledge besides the basic knowledge of language grammar. And with the abundance of languages that exist, the challenge increases exponentially.

III. RELATED WORK

While looking for a high-performance coreference tool that is easy to integrate with any NLP system with application perspective, an extensive search was done for the existing survey papers with the idea to get some clues. The interest of the search was mainly towards a comparative analysis of the tools available in terms of their performance and ease of usability. Though, with not much success in the review of tools, some of the review papers were revisited to understand their findings. This quest became the driving force behind this paper. The understanding of other reviews on coreference has been mentioned here.

The survey study of research papers about coreference resolution was done by P. Elango that focused on linguistic-based approaches and machine learning-based approaches [2]. Linguistic and domain knowledge dependent algorithms like Hobbs' algorithm, Centering theory and variations of Centering theory were studied under linguistic-based category. For machine learning-based approaches statistical methods like naive-bayes based model, decision-tree based approach, conditional random field approach and its variations, clustering approach and corpus-based approaches were studied. Zheng et al. reviewed coreference resolution methods for the clinical domain with the intention to build an end-to-end coreference system [3]. They explored the methodologies based on heuristic method, supervised approaches and unsupervised approaches. The corpus used by them for exploring coreference was ontology-based due to a specific pattern found in clinical data, which can be easily captured using an ontological approach. Beheshti et al. reviewed and analyzed the coreference resolution methods and tools for a cross-document scenario [4]. A systematic in-depth study of some of the popular coreference and named-entity resolution tools like LingPipe[3], Supersense tagger[4], AFNER[5], Stanford-NLP[6], OpenNLP[7] and Alchemy API[8] with a focus towards cross-document coreference analysis was explored in their survey. Suthankar et al. presented a review of the scope of anaphora and coreference resolution task in terms of entity resolution [1]. A detailed analysis of the dataset, evaluation metrics and algorithm used from rule-based methods to deep learning methods for the anaphora and coreference resolution is presented in this paper.

After going through the reviews done in the past it was found that either the reviews are too old to include all latest developed coreference tool or the focus is mainly from a different perspective or it is only a paper study review. This paper attempts to present a systematic study of coreference resolution tools in the light of both performance, usability and ease of integration with any NLP application. The outcome of the work is to propose a high-performance coreference tool or library that can be integrated with any NLP application system for usage.

IV. COREFERENCE TOOLS CHOSEN FOR STUDY

While exploring for a good performance coreference resolution tool or library for integration with an NLP application, both licensed as well open source coreference tools and python libraries were considered. The literature review clearly indicates that the experimentation on coreference tool started around two decades ago. Since then a long list of open source coreference resolution tools and libraries from various NLP labs have been released across the globe. At the commercial level, no tool was found listed on the web. The possible reason could be the fact that coreference resolution appears almost at the end of NLP pipeline and there are hardly any commercial organizations engaged in developing an NLP tool from end to end on a commercial basis. The focus of these organization remains only on a specific application or domain. Hence, the focus of this study have been limited to open source coreference resolution tools and libraries. Also, while exploring the tools and literature, it was found that most experimented coreference resolution was of anaphoric nature mainly pronominal type.

The coreference resolution tools and libraries considered for this study are ARKRef [5], BART [6], Berkeley Coreference Resolution tool [7], GATE coreference tool [8], Guitar [9], Illinois Coreference Package [10], JavaRAP [11, 12], OpenNLP[7], Reconcile [13], RelaxCor [14], Stanford Deterministic Coreference Resolution system [15, 21], spacy python library [16] and CorefGraph python library [17]. These were studied for the technique used for coreference resolution, input file format, output file format, type of interface, any separate preprocessing and finally the algorithm approach used for designing the tool.

Table 1 lists out the labs/organizations/individuals where the tool was developed, URLs of coreference tools and programming language used for coding, version information and licenses of the tools studied here. Also mentioned in the table is the update status of the tools. The information has been collated to get a consolidated view of all the coreference tools and libraries. The table clearly indicates that most of the tools are updated and maintained by their respective labs since their release.

[3] http://alias-i.com/lingpipe/

[4] http://sites.google.com/site/massiciara/

[5] http://afner.sourceforge.net/afner.html

[6] http://nlp.stanford.edu/

[7] http://opennlp.apache.org/

[8] http://www.alchemyapi.com/

Table 1: List of open source Coreference tools and libraries at a glance

S.No.	Name of the Tool	Developed By	Source	Language	Last Modified	Version	License
1	ARKRef	Carnegie Mellon University	http://www.cs.cmu.edu/~ark/ARKref/	Java	2013	none	GPL
2	BART	Johns Hopkins University	http://www.sfs.uni-tuebingen.de/~versley/BART/	Java	2008	1.0	GPL, Apache
3	Berkeley Entity Resolution System	University of California, Berkeley	http://nlp.cs.berkeley.edu/projects/coref.shtml	Scala	2015	1.1	GPLv3
4	GATE	The University of Sheffield	https://gate.ac.uk/sale/tao/pronom-coref	Java	2016	8.5.1	GPL
5	Guitar	University of Essex	http://cswww.essex.ac.uk/Research/nle/GuiTAR/gtarNew.html	Java	2007	3.0.3	GPL
6	Illinois Coreference Package	University of Illinois	https://cogcomp.cs.illinois.edu/page/software_view/Coref	Java	2008	1.3.2	Academic Use license
7	JavaRAP	National University of Singapore	http://aye.comp.nus.edu.sg/~qiu/NLPTools/JavaRAP.html	Java	2011	1.13	GPL
8	OpenNLP	Apache Software Foundation	https://issues.apache.org/jira/browse/OPENNLP-48	Java	2010	1.6.0	Apache v2.0
9	Reconcile	Cornell University	https://www.cs.utah.edu/nlp/reconcile/	Java	2010	1.0	GPL
10	RelaxCor	Universitat Politècnica de Catalunya	http://nlp.lsi.upc.edu/relaxcor/	Perl	2012	1.1	GPL
11	Stanford Deterministic Coreference Resolution System	Stanford University	http://nlp.stanford.edu/software/dcoref.shtml	Java	2016	3.9.2	GPL v3
12	Spacy	Explosion AI	https://spacy.io/	Python	2017	2.1.3	MIT
13	CorefGraph	Rodrigo Agerri, University of Deusto.	https://pypi.org/project/corefgraph	Python	2017	1.2.3	OSI: Apache Software License

Table 2 consolidates the information regarding the functioning of these tools. It collates the information about the input file format, output file format, type of interface available, if the tool requires any kind of separate preprocessing and the algorithm approach used for designing the tool.

From *Table 2* it is evident that different input file format supported by the tools are raw text, CoNLL and XML. Raw text format is the preferred format except for Guitar which accepts only XML input and RelaxCor that supports only CoNLL format. Similarly, the different output file format generated by the tools are coreference annotated data file, XML format and CoNLL format file. Besides these JavaRAP can produce a list of anaphora-antecedent pair format, reconcile can produce MUC-6 format, RelaxCor can produce HTML output and Stanford can also generate json and serialized format output. The interface supported by most of these tools is command line. Besides this BART, Illinious, JavaRap, Stanford and Spacy also support online and some are available as a framework, library, API *etc.* such as Spacy, CorefGraph, GATE. Since coreference resolution requires preprocessing, most tools support in-built preprocessing except ARKRef and GuiTAR where external preprocessing tools are required.

Table 2: Usage formats of the coreference tools.
(Y- yes, N- no, # - anaphora-antecedent pair, @ - more formats available, * - available but not working)

S.No.	Name	Input format			Output Format			Interface			Preprocessing Included	Approach
		Raw Text	CoNLL	Other	Tagged	CoNLL	Other	Online	Command line	Other		
1	ARKRef	Y	N	N	Y	N	Xml	N	Y	N	N	Rule based noun phrase coreference model. Uses external tools for preprocessing [5]
2	BART	Y	Y	N	Y	N	Xml	Y	Y	Library	Y	Machine learning using syntax and knowledge based feature [6]
3	Berkeley	Y	Y	N	Y	Y	N	N	Y	Library	Y	Machine learning using mention-ranking architecture [7]
4	GATE	Y	Y	N	N	N	Xml	N	N	GUI Frame Work	Y	Rule based pronominal coreference resolution approach [8]
5	Guitar	N	N	Xml	N	N	Xml	N	Y	API	N	Based on MARS approach [9]
6	Illinious	Y	Y	N	Y	Y	N	Y	Y	Library	Y	Machine learning approach enhanced with feature set and domain knowledge with constraints [10]
7	JavaRAP	Y	N	Xml	Y	N	Pair#	Y*	Y	N	Y	Based on classic resolution of anaphora procedure [11,12]
8	OpenNLP	Y	N	N	Y	N	N	N	N	Toolkit	Y	Machine learning approach that identifies only noun phrase mentions
9	Reconcile	Y	N	N	Y	N	MUC-6	N	Y	N	Y	Supervised machine learning approach using Weka toolkit [13]
10	RelaxCor	N	Y	N	N	Y	Html	N	Y	N	Y	Hyper graph partitioning approach using relaxation labeling [14]
11	Stanford	Y	Y	N	Y	Y	Json, Xml@	Y	Y	Y	Y	Deterministic technique for mention detection followed by multi-pass sieve approach for coreference resolution [15]
12	Spacy	Y	N	N	Y	N	N	Y	Y	Y	Y	Rule-based mentions-detection followed by a feed-forward neural-network to compute a coreference score for each pair of potential mentions [16]
13	CorefGraph	Y	N	N	Y	N	N	N	Y	N	N	Multilingual rule-based system based on the Multi Sieve Coreference System [17]

To understand these tools in terms of performance and ease of usage, some of these tools were chosen for a black box usage study. Section V discusses the experimental approach followed in detail.

V. Experimental Study

From the list of coreference resolution tools and libraries mentioned in *Table 1,* five are selected for further the experimental study. The criteria for choosing the tool/library for experimental study is as mentioned:

i) Should be easily accessible/downloadable
ii) Should be able to install/include with ease
iii) Not much additional software requirement
iv) Should be compatible with the recent version of the software and machines.

Along with the above criteria, the last modification time was also considered while choosing the tool as it suggests the maintainability of the tool. These criterias have been curated with the aim of selecting a coreference tool which can be integrated with any python based NLP application with ease. Based on these parameters, the tools/library considered for further experimental study is BART, Berkeley, Reconcile, Stanford and Spacy.

Table 3: Results of coreference resolution evaluation scores using CoNLL-2012 dataset

Name	MUC			B^3			CEAF-E			Average F1
	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	
BART	61.01	60.25	60.63	63.24	68.49	65.76	40.24	41.64	40.93	55.77
Berkeley	72.01	66.48	69.13	62.46	53.07	57.38	56.26	53.25	54.71	60.41
Reconcile	57.96	58.37	58.16	58.65	51.04	54.58	39.26	40.46	39.85	50.87
Stanford	75.87	67.15	71.24	69.01	55.67	61.63	62.01	55.30	58.46	63.78
Spacy	77.91	71.17	74.39	71.06	57.09	63.31	63.14	56.17	59.45	65.72

The method followed for experimentation was:

i) Install these tools on a local machine
ii) Use a standard test data for which gold standard reference data is available for evaluation.
iii) Test all five systems with the raw test set as input
iv) Test the annotated output file against the gold standard reference data
v) Evaluate the result with the standard automatic metrics used for evaluating coreference resolution links

A. Data

For the experiment here, the CoNLL 2012[9] dataset is used. This dataset was released as a part of CoNLL 2012 shared task on coreference resolution. The Based input was given in plain text format and the output generated was received in annotated file format. This output was converted to CoNLL format for evaluation.

B. Evaluation Metrics Used

Since the outcome of the coreference resolution is in the form of clusters/chains, the metric is expected to evaluate these chains against the gold standard data set. The metric used for evaluation is the standard coreference metric MUC, B^3 and CEAF-E which is discussed here as:

i) *MUC:* MUC is the most widely-used metric based on coreference links [18]. It counts the number of common links between the reference gold set and the system output. The precision is computed as the number of common links between gold chains and the system-output chains divided by the number of links in the system-output chains. The recall is computed as the number of common links between the gold chains and the system-output chains are divided by the number of links in the gold chains.
ii) *B-Cubed:* B-CUBED also referred as B^3 is a mention-based metric, i.e. the overall recall/precision is computed based on the recall/precision of the individual mentions and then averaged to obtain the overall recall and precision [19]. This metric considers singleton mentions which were missing in MUC evaluations.
iii) *CEAF-E:* A Constrained Entity Alignment F-Measure (CEAF) metric scores a coreference resolution by finding an optimal one-to-one mapping (or alignment) between the gold chains and the system-output chains [20]. It is able to address the issue of B^3 metric which uses a chain more than once in precision and recall evaluation. It evaluates by finding the best one-to-one map as a maximum bipartite matching problem. CEAF uses a similarity metric for each pair of entities i.e. a set of mentions to measure the goodness of each possible alignment.

VI. Result and Analysis

Using MUC, B^3 and CEAF-E metrics, precision and recall of the outputs are calculated and combined to give an F-score measure. The F-scores of all 3 metrics is averaged to get an integrated performance of all the tools.

Table 3 shows the experiment result of coreference resolution analysis. *Figure 2* shows the performance of each of the coreference tool studied using the standard evaluation metrics. *Figure 3* displays the F-scores of each metric for the tools into consideration. The results show that spacy library has outperformed all the other chosen systems by an improved average F-score of 1.94 percent. The reason can be attributed to the deep learning algorithm approach followed by this library. The spacy coreference resolution approach involves a rule-based mentions-detection method to identify the coreference pairs which is then scored using a feed-forward neural network model. This library requires the user to load the pre-trained neural coreference model which is available in three different sizes for English language. The library also provides the option to train the neural model for

[9] http://conll.cemantix.org/2012/data.html

other languages provided a large dataset is available.
Through this hands-on experience, the standard coreference tools were used and understood the way to include these as a part of any NLP pipeline.

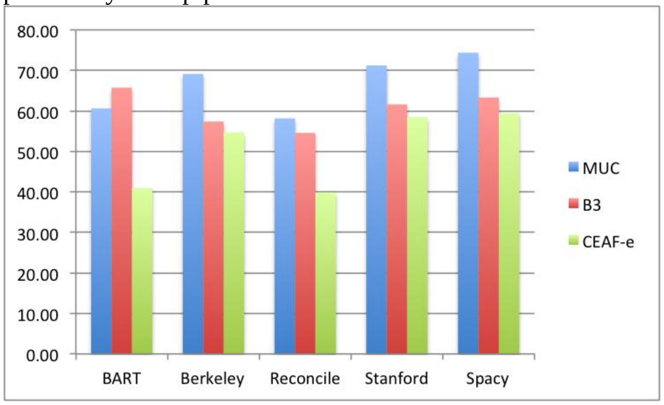

Figure 2: Performance of each coreference tool experimented

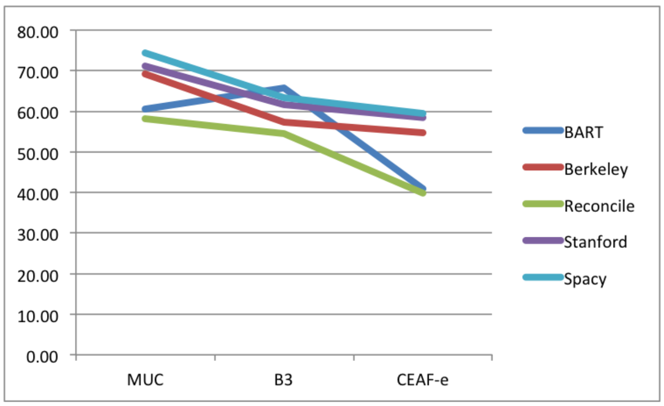

Figure 3: F-measure of the three metrics for the tools experimented

The results show that spacy library is the better performing coreference resolution system among the systems considered here for the experiment. This experiment presented a black box approach to choose the best performing coreference tool for integrating it with any NLP application.

VII. Conclusion and Future Scope

This paper presented a theoretical and comprehensive study of 13 open source coreference resolution tools and libraries on the parameters of their approach, maintainability and usage *etc.* From the tools studied BART, Berkeley, Reconcile, Stanford and Spacy tools were selected further for an experimental study and analysis. A black-box testing approach was followed for a performance study using CoNLL dataset. The output received was evaluated using the standard coreference evaluation metrics MUC, B^3 and CEAF-E. The F-scores show that Spacy outperforms all the selected tools. This study can be used as a guiding element by the research community in selecting an open source coreference resolution tool while developing any NLP application for understanding and generation. In future, based on this study, a system can be build by integrating multiple coreference tools into NLP pipeline to maximize the performance of the NLP application.

References

[1] R. Sukthanker, S. Poria, E. Cambria, R. Thirunavukarasu, *"Anaphora and Coreference Resolution: A Review."* CoRR, abs/1805.11824, **2018**.

[2] P. Elango, *"Coreference resolution: A survey."* University of Wisconsin, Madison, WI, **2005**.

[3] J. Zheng, W. W. Chapman, R. S. Crowley, G. K. Savova, *"Coreference resolution: A review of general methodologies and applications in the clinical domain".* Journal of biomedical informatics, **2011**.

[4] B. Benatallah, S. Venugopal, S. H. Ryu, H. R. Motahari-Nezhad, W. Wang, *"A systematic review and comparative analysis of cross-document coreference resolution methods and tools".* Computing, 1-37, **2016.**

[5] B. O'Connor, M. Heilman, *"ARKref: A rule-based coreference resolution system."* arXiv preprint arXiv:1310.1975 , **2013.**

[6] Y. Versley, S. P. Ponzetto, M. Poesio, V. Eidelman, A. Jern, J. Smith, A. Moschitti, *"BART: A modular toolkit for coreference resolution."* In Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies, **2008.**

[7] D. Greg, D. L. Wright Hall, D. Klein. *"Decentralized Entity-Level Modeling for Coreference Resolution."* ACL, **2013**.

[8] M. Dimitrov, K. Bontcheva, H. Cunningham, D. Maynard, *"A Lightweight Approach to Coreference Resolution for Named Entities in Text",* Anaphora Processing: Linguistic, cognitive and computational modelling 263, **2005.**

[9] M. Poesio, K. M. Alexandrov, *"A General-Purpose, Off-the-shelf Anaphora Resolution Module: Implementation and Preliminary Evaluation."* LREC, **2004.**

[10] K. W. Chang, R. Samdani, A. Rozovskaya, M. Sammons, D. Roth, *"Illinois-Coref: The UI system in the CoNLL-2012 shared task."* In Joint Conference on EMNLP and CoNLL-Shared Task (pp. 113-117). Association for Computational Linguistics, **2012.**

[11] S. Lappin, J. L. Herbert, *"An algorithm for pronominal anaphora resolution."* Computational linguistics 20.4 (1994): 535-561, **1994.**

[12] Q. Long, K. Min-Yen, C. Tat-Seng, *"A Public Reference Implementation of the RAP Anaphora Resolution Algorithm."* In proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004). Vol. I, pp. 291-294, **2004.**

[13] V. Stoyanov, C. Cardie, N. Gilbert, E. Riloff, D. Buttler, D. Hysom, *"Coreference Resolution with Reconcile",* Proceedings of the Conference of the 48th Annual Meeting of the Association for Computational Linguistics, ACL, **2010.**

[14] E. Sapena, L. Padró, J. Turmo, *"RelaxCor: A Global Relaxation Labeling Approach to Coreference Resolution."* In Proceedings of the ACL Workshop on Semantic Evaluations (SemEval- 2010), Uppsala, Sweden, July**, 2010.**

[15] H. Lee, Y. Peirsman, A. Chang, N. Chambers, M. Surdeanu, D. Jurafsky, *"Stanford's multi-pass sieve coreference resolution system at the CoNLL-2011 shared task."* In Proceedings of the fifteenth conference on computational

natural language learning: Shared task (pp. 28-34). Association for Computational Linguistics, **2011.**

[16] K. Clark, C. D. Manning, *"Deep reinforcement learning for mention-ranking coreference models. "* In Proceedings of the Conference of the Emperical Methods in Natural Language Processing, **2016.**

[17] R. Agerri, J. Bermudez, G. Rigau, *"IXA pipeline: Efficient and Ready to Use Multilingual NLP tools."* In LREC (Vol. 2014, pp. 3823-3828), **2014.**

[18] M. Vilain, J. Burger, J. Aberdeen, D. Connolly, L. Hirschman, *"A model-theoretic coreference scoring scheme."* In Proceedings of the 6th conference on Message understanding (pp. 45-52). Association for Computational Linguistics, **1995.**

[19] A. Bagga, B. Baldwin, B., *"Algorithms for scoring coreference chains."* In The first international conference on language resources and evaluation workshop on linguistics coreference (Vol. 1, pp. 563-566), **1998.**

[20] X. Luo, *"On coreference resolution performance metrics."* In Proceedings of the conference on human language technology and empirical methods in natural language processing (pp. 25-32). Association for Computational Linguistics, **2005.**

[21] K. Raghunathan, H. Lee, S. Rangarajan, N. Chambers, M. Surdeanu, D. Jurafsky, C. Manning, *"A multi-pass sieve for coreference resolution."* In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, **2010.**

[22] A.H.N. Kishore, M. Saravanan, *"An Extensive Evaluation of Anaphora Resolution Based Abstract Generation System."* International Journal on Computer Science and Engineering, 5(1), p.8, **2013.**

[23] H.M. Alfawareh, S. Jusoh, *"Resolving ambiguous entity through context knowledge and fuzzy approach. "*International journal on computer science and engineering (IJCSE), 3(1), pp.410-422, **2011.**

Authors Profile

Ms. Sandhya Singh pursued her Bachelor of Science (CS) from Banasthali University, Rajasthan and Master of Science (CS) from Banaras Hindu University, UP. She is currently pursuing her Ph.D. from Banasthali University, Rajasthan, India in the field of Natural Language Generation. Her research interest includes the field of Machine Translation, Natural Language Generation and Text Summarization.

Mr. Krishnanjan Bhattacharjee is a Principal Technical Officer at C-DAC, Pune, Maharashtra, India. He obtained his Ph.D. from Pune University in the field of Discourse Analysis. His research interest includes the areas of Big Data Analytics, Machine Learning in terms of Language model creations for training, Q-A systems, Decision Support Systems and Sentiment Analysis.

Mr. Hemant Darbari is the Director General of C-DAC, India, a premier R&D organization in Information Technologies and Electronics. He has been working in the area of Artificial Intelligence since 1985. His research interests includes the area of Natural Language Processing (NLP), Machine assisted Translation (MT), Information Extraction and Information Retrieval (IE/IR), Intelligent Language Tutoring systems, Speech Technology (Automatic Speech Recognition & Text to Speech System) and Mobile computing and Simulations.

Ms. Seema Verma received her PhD at Banasthali University. She is a Professor at the Department of Electronics, Banasthali University, Rajasthan, India. Her research interest includes issues related to communication System, wireless communication, VLSI Design, MIMO of DM, cryptography & networks security, turbo codes, LDPC codes. She is a Fellow of IETE and member of Indian Science Congress, ISTE.

