

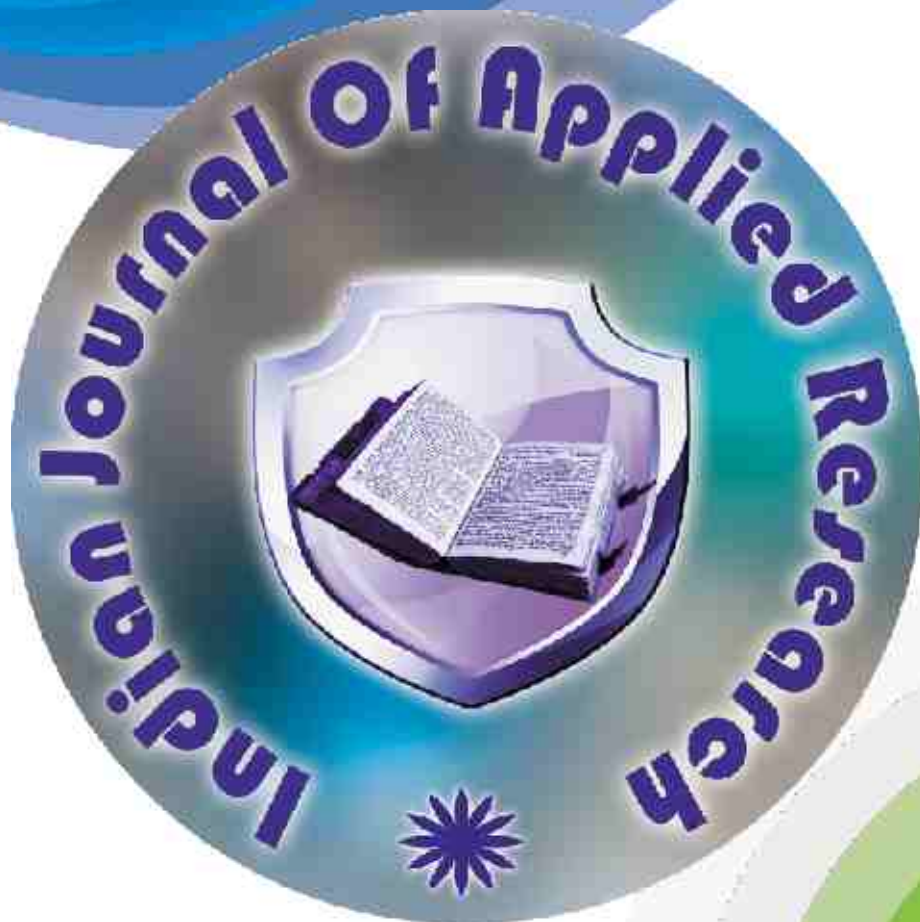
₹ 200

ISSN - 2249-555X

**Volume : 1**

**Issue : 5**

**February 2012**



**Journal for All Subjects**

[www.ijar.in](http://www.ijar.in)

Listed in International ISSN Directory, Paris.



ISSN - 2249-555X

# Indian Journal of Applied Research

## Journal for All Subjects

### Editor-In-Chief

**Dr A Kumar**

Director, College Development Council (CDC)  
Director, Internal Quality Assurance Cell (IQAC)  
Professor in Management,  
Department of Business Administration, Faculty of Management,  
Bhavnagar University,

### Editorial Advisory Board

**Dr. S. N. Pathan**  
Maharashtra

**Dr. SM. Ramasamy**  
Gandhigram

**Dr. M. M. Goel**  
Kurukshetra

**Dr. S. Ramesh**  
Tamil Nadu

**Dr Ramesh Kumar Miryala**  
Nalgonda.

**Dr. B. Rajasekaran**  
Tirunelveli

**Dr. A. R. Saravankumar**  
Tamilnadu

**Dr. Roy M. Thomas**  
Cochin

**Dr. G. Selvakumar**  
Salem

**Dr. Apurba Ratan Ghosh**  
Burdwan

**Dr. Shrawan K Sharma**  
Uttarakhand

**Dr. Sudhanshu Joshi**  
Uttarakhand

**Prof. (Dr.) B Anandampilai**  
Pudhukottai

#### Advertisement Details

Position	B/W (Single Color)	Fore Color
Full Inside Cover	₹ 6000	₹ 12500
Full Page (Inside)	₹ 5000	-

#### Subscription Details

Period	Rate	Discount	Amount Payable
One Year (12 Issues)	₹ 2400	Nil	₹ 2400
Two Year (24 issues)	₹ 4800	₹ 200	₹ 4600
Three Year (36 issues)	₹ 7200	₹ 300	₹ 6900
Five Year (60 issues)	₹ 12000	₹ 600	₹ 11400

You can download the Advertisement / Subscription Form from website [www.ijar.in](http://www.ijar.in). You will require to print the form. Please fill the form completely and send it to the **Editor, INDIAN JOURNAL OF APPLIED RESEARCH** along with the payment in the form of Demand Draft/Cheque at Par drawn in favour of **INDIAN JOURNAL OF APPLIED RESEARCH** payable at Ahmedabad.

1. Thoughts, language vision and example in published research paper are entirely of author of research paper. It is not necessary that both editor and editorial board are satisfied by the research paper. The responsibility of the matter of research paper/article is entirely of author.
2. Editing of the Indian Journal of Applied Research is processed without any remittance. The selection and publication is done after recommendations of atleast two subject expert referees.
3. In any condition if any National/International University denies accepting the research paper published in IJAR, then it is not the responsibility of Editor, Publisher and Management.
4. Only the first author is entitle to receive the copies of all co-authors
5. Before re-use of published research paper in any manner, it is compulsory to take written permission from the Editor-IJAR, unless it will be assumed as disobedience of copyright rules.
5. All the legal undertaking related to Indian Journal of Applied Research is subject to Ahmedabad Jurisdiction.
7. The research journal will be send by normal post. If the journal is not received by the author of research papers then it will not be the responsibility of the Editor and publisher. The amount for registered post should be borne by author of the research paper in case of second copy of the journal.

**Editor,**

**Indian Journal Of Applied Research**

8-A, Banans, Opp. SLU Girls College, New Congres Bhavan, Paldi,  
Ahmedabad-380006, Gujarat, INDIA

Contact.: +91-9824097643 E-mail : [editor@ijar.in](mailto:editor@ijar.in)

# Index

Sr. No	Title	Author	Subject	Page. No.
1.	Assay Of Triphenylmethane Reductase Enzyme And PCR-Based Identification Of TMR Gene In Enterobacter Asbriae Strain XJUHX-4TM	Tina Mukherjee, Mounita Bhandari, Manas Das	Biotechnology	1-2
2.	An Analysis Of Growth Of Credit Card Industry	Dr. A. Vinayagamoorthy, K. Senthikumar	Commerce	3-5
3.	Impact Of Pre-Merger And Post Merger On Financial Performance (With Reference To Private Sector Banks)	Dr. Shital Vekariya	Commerce	6-8
4.	Relativity On Climate And Competencies In Human Resource Development With Reference To Neyveli Lignite Corporation Ltd,	S. Jayakumar. Dr. R. Ramachandran	Commerce	9-11
5.	Human Resource Outsourcing: A Strategy For Gaining Competitive Advantage	Dr. Santosh M. Singh	Commerce	12-13
6.	Relationship Between EVA And ROI And MVA (A Case Study Of Ten Manufacturing Industries In India)	Dr. Shivani Gupta	Commerce	14-15
7.	Modeling The Traits Of An Effective Teacher At Higher Education	Dr. Haridayal Sharma	Commerce	16-17
8.	Mahatma Gandhi National Rural Employment Guarantee Act (Mgnrega): Issues And Challenges	Dr. Mohd. Ashraf Ali, Mushtaq Ahmad	Commerce	18-20
9.	Standardisation And Grading	Viram. J. Vala, Dr. Vijay Kumar Soni	Commerce	21-22
10.	Profitability Of Selected Information Technology Companies In India	Dr. M. Jegadeeshwaran, C. Udaya	Commerce	23-25
11.	Emerging Trends In The Indian Media And Entertainment Industry	Dr Mahalaxmi Krishnan	Commerce	26-27
12.	Inventory Management Strategies And Control Techniques: An Empirical Investigation Of Small Scale Industries	Vipul Chalotra, Neetu Andotra	Commerce	28-30
13.	A Study On Performance Indicators Of Commercial Banks	Dr. G. Ganesan, P. Parthasarathy	Commerce	31-33
14.	Improved Approaches To Coreference Resolution In Machine Learning	Kuldeep Singh Raghuwanshi, Ashwini Kumar Verma	Computer Science	34-37
15.	Security Issues & Controls In Cloud Computing	V. Naga Lakshmi	Computer Science	38-40
16.	Human Development Index Of De-Notified Nomadic Castes In Maharashtra Division: A Study Of Jalna And Aurangabad Districts	Dr. Ashok Pawar	Economics	41-43
17.	Public Private Partnership In Rural & Urban Projects In India	Dr. Ashok S. Pawar, Dr. Shankar B. Ambhore	Economics	44-45
18.	Populace Insight On Development In Public Health Sector Of India Subsequent To Functioning Of National Rural Health Mission	Krishnakant Sharma	Economics	46-49
19.	Problems Of Rural Women Entrepreneurs In India: A Conceptual Overview	C. Jeyasri Usha N Devi, Dr. A. Sankaran	Economics	50-52
20.	Poverty Of Banjara And Vanjari Communities In India	Tidke Atish S., Dr. Pawar Ashok S.	Economics	53-54
21.	India And China: Economic Reforms And WTO	Dr. Surinder Kumar Singla, Dr. Kulwinder Singh	Economics	55- 56
22.	Implementing Life Skill Education Strategies In Teaching – Learning Process	R. Kalaiselvi, Dr. A. Palanisamy, Dr. A R. Saravanakumar	Education	57-59

23.	Utilisation Of Modern Technology By The Teachers In Pupil Processing Organisation	Dr. P.Paul Devanesan, Dr A. Selvan	Education	60-61
24.	Impact Of Vocational Training On Students	K.Sudha Rani, G.Umapathi, Dr. T. Ananda,	Education	62-63
25.	A Study On Emotional Intelligence Of Secondary School Teachers	Dr. Umme Kulsum, Prathima H.P.	Education	64-66
26.	The Efficiency Of Feedback Strategy Of Homework On The Development Of 10th Grade EFL Writing Skill In Al-Karak Educational Directorate	Majid Al- Khataybeh, Areej Al-Shourafa`	noitacudE	67-74
27.	Perspectives Of Stress Management In Education System	M. Meenakshisundaram, G. P. Raja, Dr. A R. Saravanakumar	Education	75-76
28.	Attention Regulation Of Meditators And Non-Meditators Of Class IX	G. Madhavi Kanakadurga, Dr. D. Vasanta Kumari,	Education	77-78
29.	Role Of Psychoeducation In Teaching – Learning Process	Dr. A R. Saravanakumar, Dr. A. Balu, Dr. S. Subbiah	Education	79-80
30.	Microcontroller Driven RGB Led System For Tristimulus Surface Colorimetry	T. N. Ghorude, A. D. Shaligram	Electronics	81-83
31.	Pmgsy And Rural Roads Development In India: Economic, Financial And Maintenance Issues	K.C. Manjunath	Engineering	84-86
32.	Routing Packets On A Chip.	Naren V Tikare	Engineering	87-89
33.	Finding The Nearest Neighbors In Biological Databases	Er. Pankaj Bhambri, Dr. O.P. Gupta, Er. Franky Goyal	Engineering	90-92
34.	Factors Affecting The Sustainability Of The Asphalt Roads: A Case Study Of Irbid Inner Ring Road, Jordan	Eng. Nasr Ahmad Dr. Mihai Iliescu	Engineering	93-94
35.	Physical And Chemical Testing Of Compounded PVC	Sapna Dabade, Dr. Dheeraj Mandloi, Deepak Khare	Engineering	95-96
36.	Impact Of Organic Farming On Yield Of Some Common Crops- A Case Study.	Namrata D. Awandekar	Environmental Science	97
37.	Hydrogeologic Settings Of The North And South Brahmaputra Plains In Upper Assam: A Comparative Study	Dr. Uttam Goswami	Geology	98-100
38.	To Study Staffing Pattern In Rajasthan Public Healthcare Delivery System.	Dr. Ashwin G. Modi, Sushman Sharma	Healthcare	101-105
39.	Work And Health: A Situational Analysis Of Factory Workers	Dr. S. S. Vijayanchali, Dr. E. Arumuga Gandhi	Home Science	106-108
40.	Performance Of Camel Kid Hair: Acrylic Blended Yarn And Knitted Fabric	Suman Pant, Anjali Sharma	Home Science	109-110
41.	Impact Of Holistic Nutrition Education Package On Diabetes Mellitus Control In Middle Aged Women	Dr. Anjali Rajwade	Home Science	111-112
42.	Assessment Of Relationship Between Ida And Personal Hygiene, Nutritional Knowledge And Dietary Practices In Adolescent Girls	Dr. Anjali Rajwade	Home Science	113-114
43.	Employee Attrition And Retention In Private Insurance Sector– A HRM Challenge	Dr. J. Senthil Vel Murugan, S.Bala Murugan	Human Resource Management	115-117
44.	A Study On Impact Of Unionism On Industrial Relations In Manufacturing Sector	Jaya Ahuja	Industrial Relations	118-120

45.	Augmentation Of India's Foreign Exchange Reserve: An Analysis	Dr.S P.Mathiraj, Ar.Annadurai	International Business	121-123
46.	Films – A Techno Literary Art Form	Dr. Dipti Mehta	Literature	124-125
47.	Indirect Models Of Reading To Develop Descriptive Writing	Dr. K. Madhavi	Literature	126-128
48.	Ramkrishna Mishra Ke Upanaso Me Rajnetaik Chetavni	Dr. Sanjay Rathod, Dilip Jhadav	Literature	129
49.	Hindi Kavita Me Nari Jivan Ka Badla Swarup	Dr. Sanjay Rathod	Literature	130
50.	Impact Of IPL Sponsorship On Consumer Buying Behavior With Reference To Nagpur City	Chandrima Das	Management	131-135
51.	Crowd Sourcing –A New Management Mantra	Devi Premnath, Dr. C. Nateson	Management	136-137
52.	Small Scale Industries In India: An Evaluation Of Productivity In The Post-Liberalized Scenario	Dr. Gaurav Lodha,	Management	138-139
53.	Comparative Analysis Of Milk Products With Respect To Its Competitors With Special Reference To Karnataka Milk Federation (KMF) – At Dharwada City, Karnataka, India	Dr. N. Ramanjaneyalu	Management	140-143
54.	A Study On Work Stress In Women Employees In Coimbatore District	R. Maheswari, N. Brindha	Management	144-145
55.	Accounting For Carbon Credits	Dr. Gaurav Lodha	Management	146-148
56.	A Literature Review On The Relationship Between Training (As A Core Responsibility Of HRM) And Firm Performance.	Priya Sharma, Dr. S. L. Gupta	Management	149-152
57.	A Study On Agricultural Marketing Practices And Constraints With Special Reference To Paddy / Rice.	CM Maran, Dr Raja Pranmalai	Management	153-156
58.	Performance Of Share Price Of Indian Public Sector Banks And Private Sector Banks - Comparative Study	V. Prabakaran, D. Lakshmi Prabha	Management	157-158
59.	Intuitionistic Fuzzy Primary And Semiprimary Ideal	Dr. M.Palanivelrajan, S.Nandakumar	Mathematics	159-160
60.	Significance Of Umbilical Artery Velocimetry In Perinatal Outcome Of Fetuses With Intrauterine Growth Retardation.	Dr G S Shekhawat	Medical Science	161-163
61.	Large Adult Sacrococcygeal Teratoma: A Case Report And Review Of Literature.	Dr.Yavalkar Pa, Dr. Naik Am.	Medical Science	164-165
62.	Epidural Steroid In Low Back Ache	Dr. B. L. Khajotia, Dr. Neelam Meena	Medical Science	166-167
63.	A Comparative Study Of Second Trimester MTP With Use Of Vaginal Misoprostol And Extra Amniotic Instillation Of Ethacridine Lactate.	Dr. Ketaki Junnare, Dr. Sameer Darawade, Dr. Priyamvada Shah, Dr. Swati Mali.	Medical Science	168-169
64.	A Novel Surgical Approach For Treatment Of Sui –TVT Obturator Tape	Dr. Ketaki Junnare, Dr. Durga Karne, Dr Neelesh Risbud.	Medical Science	170-171
65.	Advantage Of Fallopian Tube Sperm Perfusion Over Intra-Uterine Insemination When Used In Combination With Ovarian Stimulation For The Treatment Of Unexplained Infertility.	Dr G S Shekhawat, Dr Pushpalata Naphade	Medical Science	172-175

66.	"Bilateral Sertoli-Leydig Cell Tumor In Postmenopausal Female" A Case Report	Dr. Priyamvada Shah, Dr. Ketakijunnare, Dr. DurgaKarne	Medical Science	176-178
67.	Pretreatment With Ephedrine For Prevention Of Pain Associated With Propofol Injection.	Dr. Kavita U Adate, Dr. Jyoti A. Solanki	Medical Science	179-181
68.	Does The Structured Teaching Programme Influence The Knowledge About Physical Wellbeing Of School Children? A Quasi Experimental Study.	Dr. S. Valliammal, Dr. Ramachandra, Raja Sudhakar	Nursing	182-184
69.	An Approach For Information Retrieval For Bookstores Using Formal Ontology	Sumit Jain, C.S.Bhatia	Ontology	185-187
70.	Analgesic Activity Of Anacardium Occidentale	A. Devadoss, C. Aparna, K. Parimala, D. Sukumar	Organic Chemistry	188-190
71.	Behaviourism : Science Or Metaphysics	Dr. Jatinder Kumar Sharma	Philosophy	191-193
72.	Multi-Dimensional Perspectives Of Obesity And Its Management	S. Dhanaraj, Dr. A. Palanisamy	Physical Education	194-196
73.	Refractive Index, Density, Excess Molar Volume, Excess Molar Refraction For Liquid Mixtures (Ethyl Ethanoate + Benzene Derivatives) At Different Temperatures	Sheeraz Akbar, Mahendra Kumar	Physics	197-199
74.	Refractive Indices, Densities And Excess Properties For Liquid Mixtures (Cetane + Alkanols) At Different Temperatures	Sheeraz Akbar, Mahendra Kumar	Physics	200-202
75.	Capacity Building For Effective Local Governance: Indian Perspectives	Dr. Pralhad Chengte	Political Science	203-205
76.	Psychological Well-Being: A Study Of Non-Institutionalized Aged	Dr. Pankaj S. Suvera	Psychology	206-208
77.	Women Empowerment Through N R E G S (With Reference To State Of West Bengal)	Dilip Kumar Karak	Social Sciences	209-211
78.	Effect Of Selected Yogic, Aerobic And Laughter Exercises On Blood Pressure Of High School Boys	Dr.Manjappa.P, Dr.Shivarama Reddy. M	Sports	212-216
79.	Association Study Between Lead And Copper Accumulation At Different Physiological Systems Of Goat By Application Of Canonical Correlation And Canonical Correspondence Analyses	Partha Karmakar, Debasis Mazumdar, Seema Sarkar (Mondal), Sougata Karmakar	Statistics	217-219
80.	Development Of Silver -Silica Nanocomposite For Novel Humidity Sensing Application	Surender Duhan	Technology	220-221



## Improved Approaches to Coreference Resolution in Machine Learning

\* Kuldeep Singh Raghuwanshi \*\* Ashwini Kumar Verma

\* Department of Computer Science, Suresh Gyan Vihar University, Jaipur

\*\* Department of Information Technology, Medicaps Institute of Technology & Management, Indore

### ABSTRACT

*We present a noun phrase coreference system that extends the work of Soon et al. (2001) and, to our knowledge, produces the best results to date on the MUC-6 and MUC-7 coreference resolution datasets. F-measures of 70.4 and 63.4, respectively. Improvements arise from two sources: extra-linguistic changes to the learning framework and a large-scale expansion of the feature set to include more sophisticated linguistic knowledge.*

**Keywords :** coreference; clustering; antecedent

### Introduction

Noun phrase coreference resolution refers to the problem of determining which noun phrases (NPs) refer to each real-world entity mentioned in a document. Machine learning approaches to this problem have been reasonably successful, operating primarily by recasting the problem as a classification task (e.g. Aone and Bennett (1995), McCarthy and Lehnert (1995)). Specifically, a pair of NPs is classified as co-referring or not based on constraints that are learned from an annotated corpus. A separate clustering mechanism then coordinates the possibly contradictory pairwise classifications and constructs a partition on the set of NPs [1]. Soon et al. (2001), for example, apply an NP coreference system based on decision tree induction to two standard coreference resolution data sets (MUC-6, 1995; MUC-7, 1998), achieving performance comparable to the best-performing knowledge-based coreference engines. Perhaps surprisingly, this was accomplished in a decidedly knowledge-lean manner: the learning algorithm has access to just 12 surface-level features.

This paper presents an NP coreference system that investigates two types of extensions to the Soon et al. corpus-based approach. First, we propose and evaluate three extra-linguistic modifications to the machine learning framework, which together provide substantial and statistically significant gains in coreference resolution precision. Second, in an attempt to understand whether incorporating additional knowledge can improve the performance of a corpus-based coreference resolution system, we expand the Soon et al. feature set from 12 features to an arguably deeper set of 53. We propose additional lexical, semantic, and knowledge-based features; most notably, however, we propose 26 additional grammatical features that include a variety of linguistic constraints and preferences. Although the use of similar knowledge sources has been explored in the context of both pronoun resolution (e.g. Lappin and Leass (1994)) and NP coreference resolution (e.g. Grishman (1995), Lin (1995)), most previous work treats linguistic constraints as broadly and unconditionally applicable hard constraints. Because sources of linguistic information in a learning-based system are represented as features, we can, in contrast, incorporate them selectively rather than as universal hard constraints.

Our results using an expanded feature set are mixed. First, we find that performance drops significantly when using the full feature set, even though the learning algorithms investigated have built-in feature selection mechanisms. We demonstrate empirically that the degradation in performance can be attributed, at least in part, to poor performance on common noun resolution. A manually selected subset of 2226 features, however, is shown to provide significant gains in performance when chosen specifically to improve precision on common noun resolution. Overall, the learning framework and linguistic knowledge source modifications boost performance of Soon's learning-based coreference resolution approach from an F-measure of 62.6 to 70.4, and from 60.4 to 63.4 for the MUC-6 and MUC-7[1] data sets, respectively. To our knowledge, these are the best results reported to date on these data sets for the full NP coreference problem.

The rest of the paper is organized as follows. In sections 2 and 3, we present the baseline coreference system and explore extra-linguistic modifications to the machine learning framework. Section 4 describes and evaluates the expanded feature set. We conclude with related and future work in Section 5.

### The Baseline Coreference System

Our baseline coreference system attempts to duplicate both the approach and the knowledge sources employed in Soon et al. (2001). More specifically, it employs the standard combination of classification and clustering described above. Building an NP coreference classifier. We use the C4.5 decision tree induction system (Quinlan, 1993) to train a classifier that, given a description of two NPs in a document,  $NP_i$  and  $NP_j$ , decides whether or not they are coreferent. Each training instance represents the two NPs under consideration and consists of the 12 Soon et al. features called Linguistically, the features can be divided into four groups: lexical, grammatical, semantic, and positional. The classification associated with a training instance is one of COREFERENT or NOT COREFERENT depending on whether the NPs co-refer in the associated training text. We follow the procedure employed in Soon et al. to create the training data: we rely on coreference chains from the MUC answerers to create (1) a *positive instance* for each anaphoric noun phrase,  $NP_j$  and its closest preceding antecedent,  $NP_i$ ; and (2) a *negative instance* for  $NP_j$  paired with each of the intervening

NPs,  $NP_{i+1}$   $NP_{i+2}$ , .. ,  $NP_{j-1}$  . This method of negative instance selection is further described in Soon et al. (2001); it is designed to operate in conjunction with their method for creating coreference chains, which is explained next.

#### Applying the classifier to create coreference chains.

After training, the decision tree is used by a clustering algorithm to impose a partitioning on all NPs in the test texts, creating one cluster for each set of coreferent NPs. As in Soon et al., texts are processed from left to right. Each NP encountered,  $NP_j$  is compared in turn to each preceding NP,  $NP_i$ , from right to left. For each pair, a test instance is created as during training and is presented to the coreference classifier, which returns a number between 0 and 1 that indicates the likelihood that the two NPs are coreferent.<sup>3</sup> NP pairs with class values above 0.5 are considered COREFERENT; otherwise the pair is considered NOT COREFERENT.[4] The process terminates as soon as an antecedent is found for  $NP_j$  or the beginning of the text is reached.

#### Baseline Experiments

We evaluate the Duplicated Soon Baseline system using the standard MUC-6 (1995) and MUC-7 (1998) coreference corpora, training the coreference classifier on the 30 "dry run" texts, and applying the coreference resolution algorithm on the 2030 "formal evaluation" texts. The MUC-6 corpus produces a training set of 26455 instances (5.4% positive) from 4381 NPs and a test set of 28443 instances (5.2% positive) from 4565 NPs. For the MUC-7 corpus, we obtain a training set of 35895 instances (4.4% positive) from 5270 NPs and a test set of 22699 instances (3.9% positive) from 3558 NPs. In results of Duplicated Soon Baseline where performance is reported in terms of recall, precision, and F-measure using the modeltheoretic MUCscoring program (Vilain et al., 1995).

The system achieves an F-measure of 66.3 and 61.2 on the MUC-6 and MUC-7 data sets, respectively. Similar, but slightly worse performance was obtained using RIPPER (Cohen, 1995), an information-gain-based rule learning system. Both sets of results are at least as strong as the original Soon results, indicating indirectly that our Baseline system is a reasonable duplication of that system.<sup>4</sup> In addition, the trees produced by Soon and by our Duplicated Soon Baseline are essentially the same, differing only in two places where the Baseline system imposes additional conditions on coreference.

The primary reason for improvements over the original Soon system for the MUC-6 data set appears to be our higher upper bound on recall (93.8% vs. 89.9%), due to better identification of NPs. For MUC-7[1], our improvement stems from increases in precision, presumably due to more accurate feature value computation.

#### Modifications to the Machine Learning Framework

This section studies the effect of three changes to the general machine learning framework employed by Soon et al. with the goal of improving precision in the resulting coreference resolution systems.

**Best-first clustering.** Rather than a right-to-left search from each anaphoric NP for the first coreferent NP, we hypothesized that a right-to-left search for a *highly likely antecedent* might offer more precise, if not generally better coreference chains. As a result, we modify the coreference clustering algorithm to select as the antecedent of  $NP_j$  the NP with the highest coreference likelihood value from among preceding NPs with coreference class values above 0.5.

**Training set creation.** For the proposed best-first clustering to be successful, however, a different method for training instance selection would be needed: rather than generate a positive training example for each anaphoric NP and its closest antecedent, we instead generate a positive training examples for its most confident antecedent. More specifically, for a non-pronominal NP, we assume that the most confident antecedent is the closest non-pronominal preceding antecedent[8]. For pronouns, we assume that the most

confident antecedent is simply its closest preceding antecedent. Negative examples are generated as in the Baseline system.

**String match feature.** Soon's string match feature (SOON-STR) tests whether the two NPs under consideration are the same string after removing determiners from each. We hypothesized, however, that splitting this feature into several primitive features, depending on the type of NP, might give the learning algorithm additional flexibility in creating coreference rules[2]. Exact string match is likely to be a better coreference predictor for proper names than it is for pronouns, for example. Specifically, we replace the SOON-STR feature with three features PRO-STR, PN-STR, and WORDS-STR which restrict the application of string matching to pronouns, proper names, and non-pronominal NPs, respectively. Although similar feature splits might have been considered for other features (e.g. GENDER and NUMBER), only the string match feature was tested here.

**Results and Discussion:** Results are based on the learning framework modifications. When used in combination, the modifications consistently provide statistically significant gains in precision over the Baseline system without any loss in recall. As a result, we observe reasonable increases in F-measure for both classifiers and both data sets. When using RIPPER, for example, performance increases from 64.3 to 67.2 for the MUC-6 data set and from 60.8 to 63.2 for MUC-7. Similar, but weaker, effects occur when applying each of the learning framework modifications to the Baseline system in isolation.

Our results provide direct evidence for the claim (Mitkov, 1997) that the extra-linguistic strategies employed to combine the available linguistic knowledge sources play an important role in computational approaches to coreference resolution. In particular, our results suggest that additional performance gains might be obtained by further investigating the interaction between training instance selection, feature selection, and the coreference clustering algorithm.

#### NP Coreference Using Many Features

This section describes the second major extension to the Soon approach investigated here: we explore the effect of including 41 additional, potentially useful knowledge sources for the coreference resolution classified. The features were not derived empirically from the corpus, but were based on common-sense knowledge and linguistic intuitions regarding coreference.[5] Specifically, we increase the number of lexical features to nine to allow more complex NP string matching operations. In addition, we include four new semantic features to allow finer-grained semantic compatibility tests. We test for ancestor-descendent relationships in Word-Net (SUBCLASS), for example, and also measure the WordNet graph-traversal distance (WNDIST) between  $NP_j$  and  $NP_i$ . Furthermore, we add a new positional feature that measures the distance in terms of the number of paragraphs (PARANUM) between the two NPs.

The most substantial changes to the feature set, however, occur for grammatical features: we add 26 new features to allow the acquisition of more sophisticated syntactic coreference resolution rules. Four features simply determine NP type, e.g. are both NPs definite, or pronouns, or part of a quoted string? These features allow other tests to be conditioned on the types of NPs being compared. Similarly, three new features determine the grammatical role of one or both of the NPs. Currently, only tests for clausal subjects are made. Next, eight features encode traditional linguistic (hard) constraints on coreference.[4] For example, coreferent NPs must agree both in gender and number (AGREEMENT); cannot SPAN one another (e.g. "government" and "government officials"); and cannot violate the BINDING constraints. Still other grammatical features encode general linguistic preferences either for or against coreference. For example, an indefinite NP (that is not in opposition to an anaphoric NP) is not likely to be coreferent with any NP that precedes it (ARTICLE).[5]



The last subset of grammatical features encodes slightly more complex, but generally non-linguistic heuristics. For instance, the CONTAINS PN feature effectively disallows coreference between NPs that contain distinct proper names but are not themselves proper names e.g. "IBM executives" and "Microsoft executives"). Two final features make use of an in-house naive pronoun resolution algorithm (PRO-RESOLVE) and a rule-based coreference resolution system (RULE-RESOLVE), each of which relies on the original and expanded feature sets described above.

**Discussion:** Discussion using the expanded feature also incorporates the learning framework changes. In comparison, we see statistically significant increases in recall, but much larger decreases in precision. As a result, F-measure drops precipitously for both learning algorithms and both data sets. A closer examination of the results indicates very poor precision on common nouns in comparison to that of pronouns and proper nouns. In particular, the classifiers acquire a number of low precision rules for common noun resolution, presumably because the current feature set is insufficient. For instance, a rule induced by RIPPER classifies two NPs as coreferent if the first NP is a proper name, the second NP is a definite NP in the subject position, and the two NPs have the same semantic class and are at most one sentence apart from each other. In comparison, the Baseline system obtains much better precision on common nouns (i.e. 53.3 for MUC-6/RIPPER and 61.0 for MUC-7/RIPPER with lower recall in both cases) where the primary mechanism employed by the classifiers for common noun resolution is its high-precision string matching facility. Our results also suggest that data fragmentation is likely to have contributed to the drop in performance (i.e. we increased the number of features without increasing the size of the training set). For example, the decision tree induced from the MUC-6 data set using the Soon feature set (Learning Framework results) has 16 leaves, each of which contains 1728 instances on average; the tree induced from the same data set using all of the 53 features, on the other hand, has 86 leaves with an average of 322 instances per leaf.

**Hand-selected feature sets:** As a result, we next evaluate a version of the system that employs manual feature selection: for each classifier/data set combination, we discard features used primarily to induce low-precision rules for common noun resolution and re-train the coreference classifier using the reduced feature set.[6] Here, feature selection does not depend on a separate development corpus and is guided solely by inspection of the features associated with low-precision rules induced from the training data. In current work, we are automating this feature selection process, which currently employs a fair amount of user discretion, e.g. to determine a precision cut-off.

In general, we hypothesized that the hand selected features would reclaim precision, hopefully without losing recall. For the most part, the experimental results support this hypothesis. In comparison to the All Features version, we see statistically significant gains in precision and statistically significant, but much smaller, drops in recall, producing systems with better F-measure scores. In addition, precision on common nouns rises substantially, as expected.[7] Unfortunately, the hand-selected features precipitate a large drop in precision for pronoun resolution for the MUC-7/C4.5 data set. Additional analysis is required to determine the reason for this.

Moreover, the Hand-selected Features produce the highest scores posted to date for both the MUC- 6 and MUC-7 data sets: F-measure increases w.r.t. the Baseline system from 64.3 to 70.4 for MUC- 6/RIPPER, and from 61.2 to 63.4 for MUC-7/C4.5. In one Variation (MUC-7/RIPPER), however, the Hand-selected Features slightly underperforms the Learning Framework modifications (F-measure of 63.1 vs. 63.2) although changes in recall and precision are not statistically significant. Overall,

our results indicate that pronoun and especially common noun resolution remain important challenges for coreference resolution systems. Somewhat disappointingly, only four of the new grammatical features corresponding to linguistic constraints and preferences are selected by the symbolic learning algorithms investigated: AGREEMENT, ANIMACY, BINDING, and MAXIMALNP.

**Discussion:** In an attempt to gain additional insight into the difference in performance between our system and the original Soon system, we compare the decision tree induced by each for the MUC-6.[8]

ALIAS = C: + (347.0/23.8)

```

ALIAS = I:
| SOON_STR_NONPRO = C:
|| ANIMACY = NA: - (4.0/2.2)
|| ANIMACY = I: + (0.0)
|| ANIMACY = C: + (259.0/45.8)
| SOON_STR_NONPRO = I:
|| PRO_STR = C: + (39.0/2.6)
|| PRO_STR = I:
||| PRO_RESOLVE = C:
|||| EMBEDDED_1 = Y: - (7.0/3.4)
|||| EMBEDDED_1 = N:
||||| PRONOUN_1 = Y:
||||| ANIMACY = NA: - (6.0/2.3)
||||| ANIMACY = I: - (1.0/0.8)
||||| ANIMACY = C: + (10.0/3.5)
||||| PRONOUN_1 = N:
||||| MAXIMALNP = C: + (108.0/18.2)
||||| MAXIMALNP = I:
||||| WNCLASS = NA: - (5.0/1.2)
||||| WNCLASS = I: + (0.0)
||||| WNCLASS = C: + (12.0/3.6)
|||| PRO_RESOLVE = I:
||| APPOSITIVE = I: - (26806.0/713.8)
||| APPOSITIVE = C:
|||| GENDER = NA: + (28.0/2.6)
|||| GENDER = I: + (5.0/3.2)
|||| GENDER = C: - (17.0/3.7)

```

Figure 1: Decision Tree using the Hand-selected feature set on the MUC-6 data set.

data set. For our system, we use the tree induced on the hand-selected features (Figure 1). The two trees are fairly different. In particular, our tree makes use of many of the features that are not present in the original Soon feature set. The root feature for Soon, for example, is the general string match feature (SOON-STR); splitting the SOON-STR feature into three primitive features promotes the ALIAS feature to the root of our tree, on the other hand. In addition, given two non-pronominal, matching NPs (SOON-STR-NONPRO=C), our tree requires an additional test on ANIMACY before considering the two NPs coreferent; the Soon tree instead determines two NPs to be coreferent as long as they are the same string. Pronoun resolution is also performed quite differently by the two trees, although both consider two pronouns coreferent when their strings match. Finally, intersentential and intrasentential pronominal references are possible in our system while intersentential pronominal references are largely prohibited by the Soon system.

### Conclusion

We investigate two methods to improve existing machine learning approaches to the problem of noun phrase coreference resolution. First, we propose three extra-linguistic modifications to the machine learning framework, which together consistently produce statistically significant gains in precision and corresponding increases in F-measure.

Our results indicate that coreference resolution systems can improve by effectively exploiting the interaction between the classification algorithm, training instance selection, and the clustering algorithm. We plan to continue investigations along these lines, developing, for example, a true best-first clustering coreference framework and exploring a "supervised clustering" approach to the problem. In addition, we provide the learning algorithms with many additional linguistic knowledge sources for coreference resolution. Unfortunately, we find that performance drops significantly when using the full feature set; we attribute this, at least in part, to the system's poor performance on common noun resolution and to data fragmentation problems that arise with the larger feature set. Manual feature selection, with an eye toward eliminating low-precision rules for common noun resolution, is shown to reliably improve performance over the full feature set and produces the best results to date on the MUC-6 and MUC-7 coreference data sets (F-measures of 70.4 and 63.4, respectively). [4] Nevertheless, there is substantial room for improvement. As noted above, for example, it is important to automate the precision-oriented feature selection procedure as well as to investigate other methods for feature selection. We also plan to investigate previous work on common noun phrase interpretation (e.g. Sidner (1979), Harabagiu et al. (2001)) as a means of improving common noun phrase resolution, which remains a challenge for state-of-the-art coreference resolution systems.

## REFERENCES

- MUC-6. 1995. Proceedings of the Sixth Message Understanding Conference (MUC-6). Morgan Kaufmann, San Francisco, CA. | R. Quinlan. 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann, San Mateo, CA. | W. Cohen. 1995. Fast Effective Rule Induction In Proceedings of the Twelfth International Conference on Machine Learning. | S. Lappin and H. Leass. 1994. An Algorithm for Pronominal Anaphora Resolution. Computational Linguistics, 20(4):535-562 | C. Sidner. 1979. Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse. PhD Thesis, Massachusetts Institute of Technology. | M. Vilain, J. Burger, J. Aberdeen, D. Connolly, and L. Hirschman. 1995. A model-theoretic coreference scoring scheme. In Proceedings of the Sixth Message Understanding Conference (MUC-6), pages 455-462, San Francisco, CA. Morgan Kaufmann. | MUC-7. 1998. Proceedings of the Seventh Message Understanding Conference (MUC-7). Morgan Kaufmann, San Francisco, CA. | D. Lin. 1995. University of Manitoba: Description of the PIE System as Used for MUC-6.



**Sara Publishing Academy**  
Indian Journal Of Applied Research  
Journal for All Subjects



**Editor,**  
**Indian Journal Of Applied Research**  
8-A, Banans, Opp. SLU Girls College,  
New Congres Bhavan, Paldi, Ahmedabad-380006.  
Contact.: +91-9824097643 E-mail : editor@ijar.in

Printed at Unique Offset, Novatsing Rupam Estate, Opp. Abhay Estate, Tavdipura, Shahibaug, Ahmedabad