

JU_CSE_GREC10: NAMED ENTITY GENERATION AT GREC 2010

Amitava Das, Tanik Saikh, Tapabrata Mondal, Sivaji
Bandyopadhyay

Department of Computer Science and Engineering
Jadavpur University,
Kolkata-700032, India

OVERVIEW

This paper presents the experiments carried out at **Jadavpur University** as part of the participation in the GREC Named Entity Generation Challenge 2010. The **Baseline system** is based on the SEMCAT, SYNCAT and SYN-FUNC features of REF and REG08-TYPE and CASE features of REFEX elements. The **Discourse level system** is based on the additional positional features: paragraph number, sentence number, word position in the sentence and mention number of a particular named entity in the document. The inclusion of discourse level features has improved the performance of the system.

BASELINE SYSTEM

The Baseline system is based on the following linguistic features:

- **REF elements**
 - SEMCAT (Semantic Category)
 - SYNCAT (Syntactic Category)
 - SYNFUNC (Syntactic Function)
- **REFEX elements**
 - REG08-TYPE (Entity type)
 - CASE (Case marker)

PERFORMANCE OF THE BASELINE SYSTEM

	Chefs	Composers	Inventors
Precision	0.63	0.68	0.70
Recall	0.69	0.60	0.64
F-Measure	0.66	0.64	0.68

DISCOURSE LEVEL SYSTEM

The Discourse system is based on the following linguistic features:

- Positional Information
 - Paragraph number
 - Sentence number
 - Word's particular position in a sentence
- REFEX
 - SEMCAT (Semantic Category)
 - SYNCAT (Syntactic Category)
 - SYNFUNC (Syntactic Function)

ANALYSIS OF NAME TYPE

1	<REFEX ENTITY="0" REG08-TYPE="name" CASE="genitive">Alain Senderens's</REFEX>	CNG
2	<REFEX ENTITY="0" REG08-TYPE="name" CASE="genitive">Senderens's</REFEX>	LNG
3	<REFEX ENTITY="0" REG08-TYPE="name" CASE="plain">Alain Senderens</REFEX>	CN
4	<REFEX ENTITY="0" REG08-TYPE="name" CASE="plain">Senderens</REFEX>	LN

EXPERIMENTAL RESULT

	Chefs			Composers			Inventors		
	P	R	F	P	R	F	P	R	F
Name	0.69	0.74	0.71	0.78	0.61	0.69	0.77	0.67	0.71
Pronoun	0.81	0.76	0.79	0.70	0.84	0.76	0.76	0.87	0.81
Common	0.76	0.87	0.81	0.37	0.44	0.40	0.44	0.65	0.68
Empty	0.92	0.88	0.90	0.86	0.92	0.89	0.72	0.65	0.68

EXPERIMENTAL RESULT

	Reg08 Type		String Accuracy	BLEU				NIST	String Edit Distance	
	Preci sion	Recall		1	2	3	4		Mean	Mean Normaliz ed
Chefs	0.66	0.70	0.57	0.68	0.70	0.76	0.81	3.70	0.77	0.38
Comp osers	0.63	0.67	0.56	0.61	0.57	0.54	0.50	3.34	1.07	0.40
Invent ors	0.60	0.62	0.50	0.55	0.54	0.52	0.49	2.90	1.25	0.47
Total	0.63	0.66	0.54	0.61	0.58	0.57	0.55	3.83	1.03	0.42