



Extraction of biomedical events using case-based reasoning

Mariana Lara Neves (mlara@cnb.csic.es)

José-María Carazo, Alberto Pascual-Montano

BioNLP'09 Event Extraction Shared Task

June 5th, 2009

Event Extraction



Interferon-gamma potentiates the antiviral activity and the expression of interferon-stimulated genes induced by interferon-alpha in U937 cells.

[Improta T](#), [Pine R](#), [Pfeffer LM](#).

Rockefeller University, New York, NY 10021.

Binding of type I interferon (IFN-alpha/beta) to specific receptors results in the rapid transcriptional activation, independent of protein synthesis, of IFN-alpha-stimulated genes (ISGs) in human fibroblasts and HeLa and Daudi cell lines. The binding of ISGF3 (IFN-stimulated gene factor 3) to the conserved IFN-stimulated response element (ISRE) results in transcriptional activation. This factor is composed of a DNA-binding protein (ISGF3 gamma), which normally is present in the cytoplasm, and other IFN-alpha-activated proteins which preexist as latent cytoplasmic precursors (ISGF3 alpha). We have found that ISG expression in the monocytic U937 cell line differs from most cell lines previously examined. U937 cells express both type I and type II IFN receptors, but only IFN-alpha is capable of inducing antiviral protection in these cells. Pretreatment with IFN-gamma potentiates the IFN-alpha-induced protection, but IFN-gamma alone does not have any antiviral activity. ISG15 mRNA accumulation in U937 cells is not detectable before 6 h of IFN-alpha treatment, peaks at 24 h, and requires protein synthesis. Although IFN-gamma alone does not induce ISG expression, IFN-gamma pretreatment markedly increases and hastens ISG expression and transcriptional induction. Nuclear extracts assayed for the presence of ISRE binding factors by electrophoretic mobility shift assays show that ISGF3 is induced by IFN-alpha within 6 h from undetectable basal levels in untreated U937 cells. Activation of ISGF3 alpha, the latent component of ISGF3, occurs rapidly. However, the increase in ISGF3 activity ultimately correlates with the accumulation of ISGF3 gamma induced by IFN-alpha or IFN-gamma. (ABSTRACT TRUNCATED AT 250 WORDS)

PMID: 1315834 [PubMed - indexed for MEDLINE]

Protein
(given)

Entity
(tagged)

Event
(tagged)

Arguments
(manual rules)

Event Extraction (approach)



Tagging entities and events:
Case-based reasoning

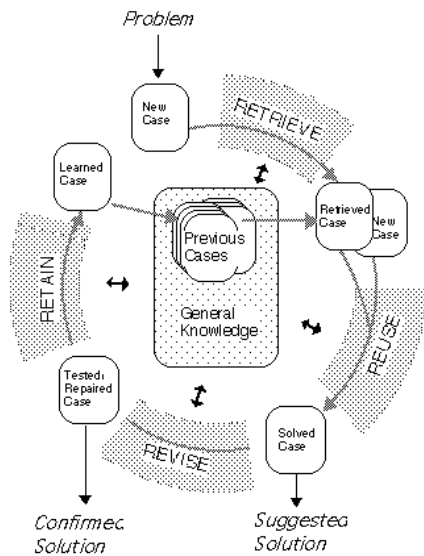
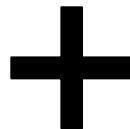
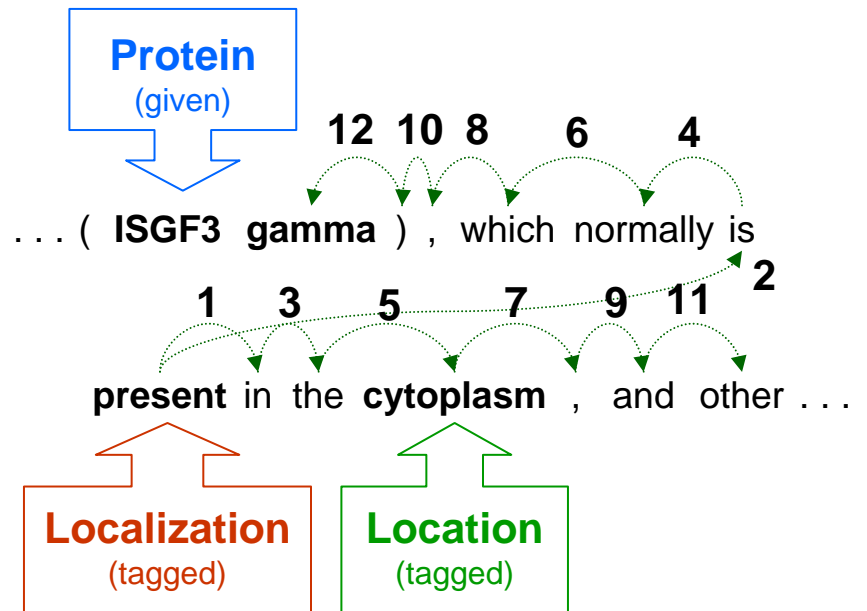


Figure 1. The CBR Cycle



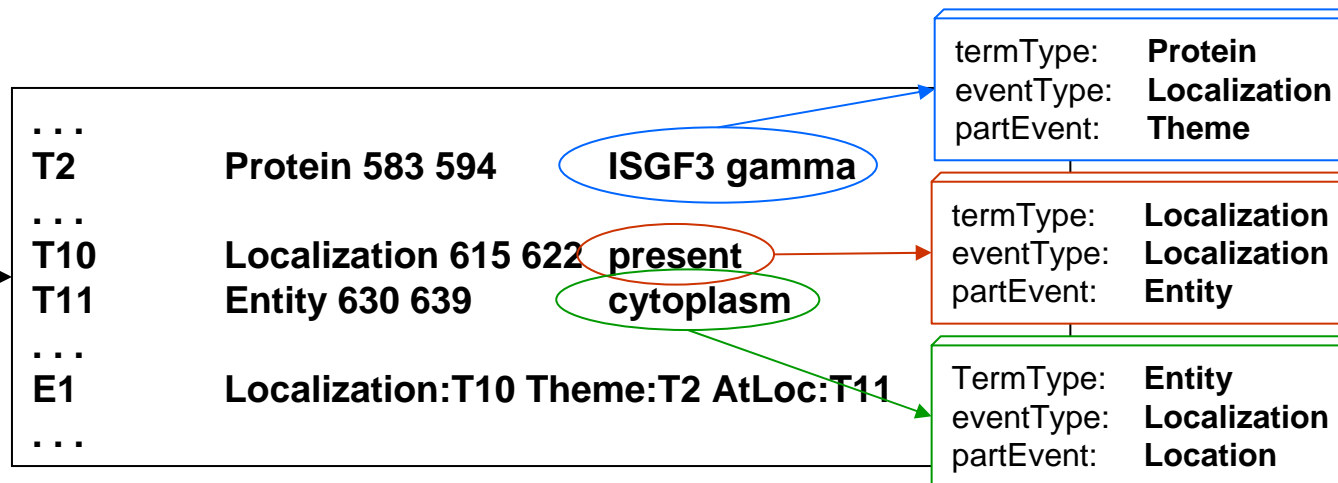
Mapping to arguments:
Manual rules



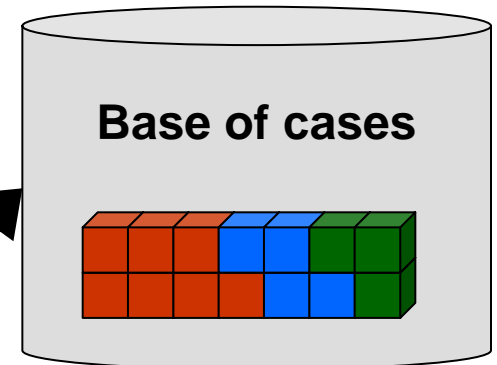


Features / Tokens	Training		Testing	
	-1	0	-1	0
stem	✓	✓	✓	✓
shape		✓		✓
posTag	✓	✓	✓	✓
chunkTag				
entityTag	✓	✓	✓	✓
termType	✓	✓	✓	?
eventType	✓	✓	✓	?
partEvent	✓	✓	✓	?

Dorsal	Aa
Bmp4	Aa1
the	p
cGKI(alpha)	aAAA(g)
patterning	pat\$a
activity	a\$vity



Tagging (training - retaining cases)



Frequency: 2

Token (-1)		Token (0)	
stem:	is	stem:	present
posTag:	VBZ	shape:	pre\$a
entityTag:	-	posTag:	JJ
termType:	-	entityTag:	-
eventType:	-	termType:	Gene_expression
partEvent:	-	eventType:	Gene_expression
		partEvent:	eventEntity

Frequency: 4

Token (-1)		Token (0)	
stem:	the	stem:	cytoplasm
posTag:	DT	shape:	a\$asm
entityTag:	-	posTag:	NN
termType:	-	entityTag:	-
eventType:	-	termType:	Entity
partEvent:	-	eventType:	Localization
		partEvent:	eventLocation

Frequency: 1

Token (-1)		Token (0)	
stem:	.	stem:	isg15
posTag:	.	shape:	AAA11
entityTag:	-	posTag:	NN
termType:	-	entityTag:	RNA
eventType:	-	termType:	Protein
partEvent:	-	eventType:	Pos_regulation
		partEvent:	eventTheme

... expression of interferon ...

... factors result in the ... of IFN-alpha-stimulated ... the binding of ISG ...

stimulated gene factor 3) to the conserved IFN-stimulated response element ... in transcriptional activation. This factor is composed of a DNA-binding protein [gamma], which normally is present in the cytoplasm, and other IFN-alpha-activated proteins which preexist as latent cytoplasmic precursors (ISGF3 alpha). We found that ISG expression in the monocytic U937 cell line differs from most cell lines examined. U937 cells express both type I and type II IFN receptors, but only capable of inducing antiviral protection in these cells. Pretreatment with IFN-gamma potentiates the IFN-alpha-induced antiviral activity. ISG15 mRNA levels peak after IFN-alpha treatment, but alone does not induce ISG expression. Pretreatment with IFN-gamma hastens ISG expression and increases its levels in the presence of ISRE binding factor 3. ISG expression is induced by IFN-alpha with or without IFN-gamma. Activation of ISGF3 alpha, the transcription factor for ISG15, is increased by the increase in ISGF3 activity induced by IFN-alpha or IFN-gamma.

PMID: 1315834 [PubMed - indexed for MEDLINE]

Tagging (retrieving cases)



Costimulation requirement for AP-1 and NF-kappa B transcription factor activation in T cells.

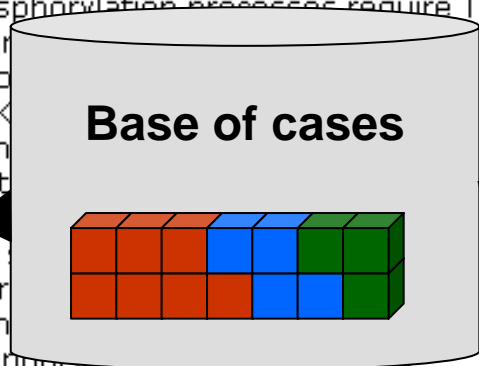
[Jung S](#), [Yaron A](#), [Alkalay I](#), [Hatzubai A](#), [Avraham A](#), [Ben-Neriah Y](#).

Lautenberg Center for General and Tumor Immunology, Hebrew University, Hadassah Medical School, Jerusalem, Israel.

The **transcriptional activity** of the **IL-2 promoter** **requires** T-cell costimulation delivered by the TCR and the auxiliary receptor **CD28**. Several transcription factors **participate in IL-2 promoter activation**, among which are AP-1-like factors and NF-kappa B. Protein phosphorylation has an important role in the regulation of these two factors: (1) it induces the transactivating capacity of the AP-1 protein c-Jun; and (2) it is involved in the release of the cytoplasmic inhibitor, I kappa B, from NF-kappa B, allowing translocation of the latter into the nucleus. We have recently shown that both phosphorylation processes require T-cell costimulation. Furthermore, in activated T cells, the kinase events are essentially similar. According to our results, however, the two processes are distinct entities. Whereas TPCK phosphorylates c-Jun and, consequently, activation of NF-kappa B, it does not phosphorylate I kappa B and, consequently, activation of NF-kappa B, it does not phosphorylate I kappa B. JNK, the MAP kinase-related kinase that phosphorylates

termType: **Positive_reg**
 eventType: **Positive_reg**
 partEvent: **Entity**

Token (-1)	Token (0)
stem: promot	stem: activ
posTag: NN	shape: a\$ion
entityTag: DNA	posTag: NN
termType: Entity	entityTag: -
eventType: Positive_reg	termType: ?
partEvent: Site	eventType: ?
	partEvent: ?



“stem” feature mandatory
 minimum matching features (MMF)
 minimum frequency of a case (MFC)

Tagging (parameters)

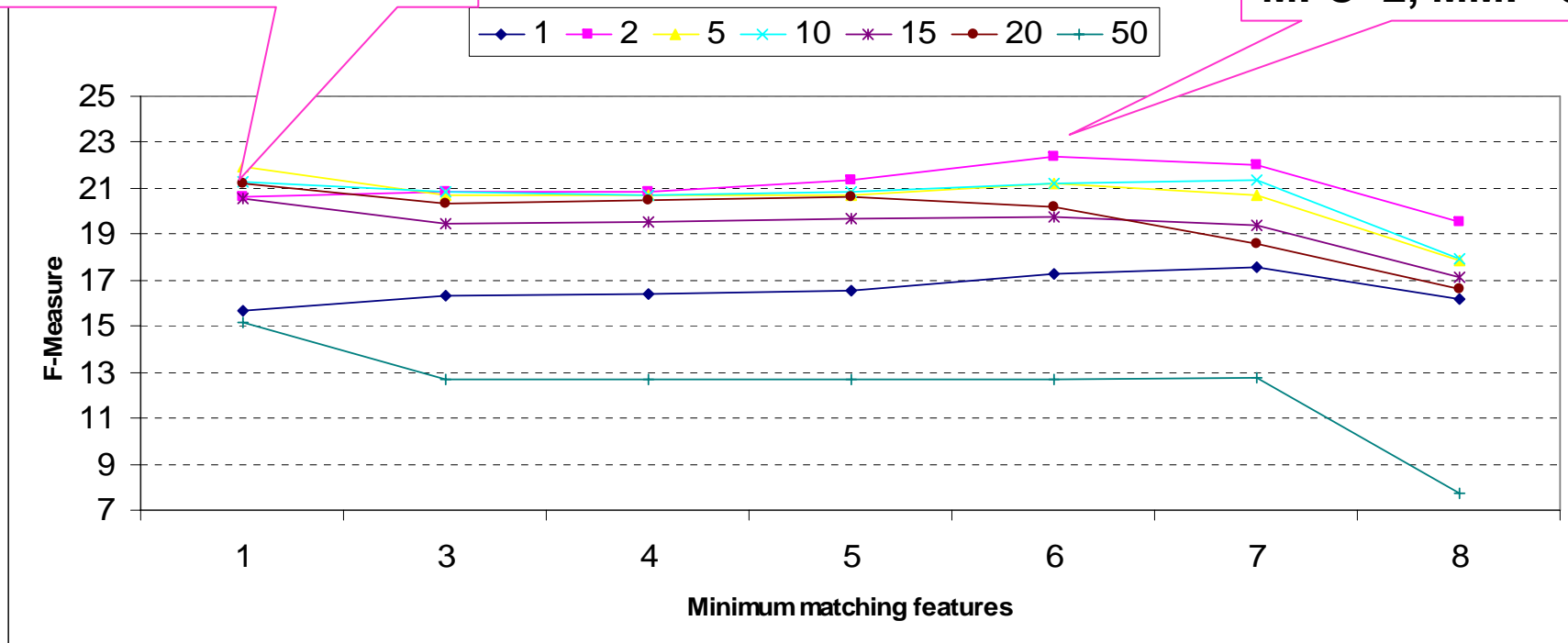


minimum matching features (MMF): 1, 3, 4, 5, 6, 7, 8

minimum frequency of a case (MFC): 1, 2, 5, 10, 15, 20, 50

(f2m1) **SUBMITTED**
MFC=2, MMF=1

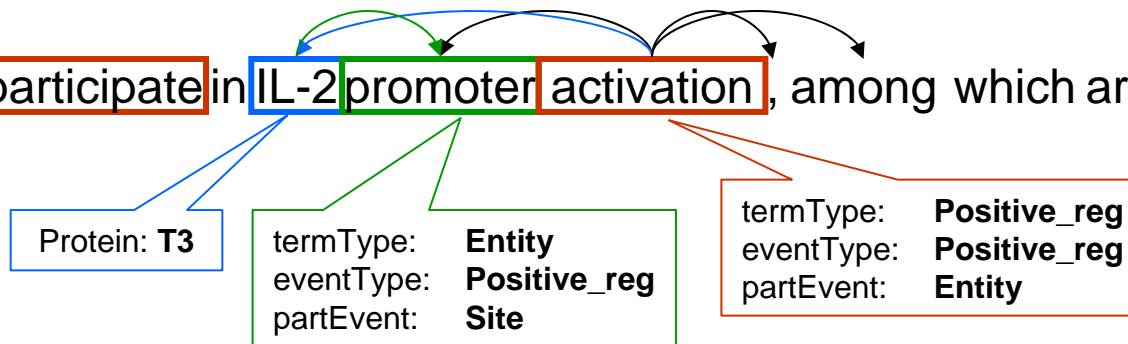
(f2m6)
MFC=2, MMF=6



Mapping (searching)



... factors participate in IL-2 promoter activation, among which are AP-1-like factors ...



...				
T11	Entity	279 287	promoter	
T12	Positive_regulation	288 298		activation
...				
E4	Positive_regulation:T12	Theme:T3	Site:T11	
...				

Mapping (manual rules)



Arguments	eventType (case-solution)	termType (token case)	partEvent (token case)	# tokens (each direction)
THEME	All	Proteins or Events	Any	20
THEME2	Binding	Proteins	Any	10 (from theme)
CAUSE	Regulation Pos. Reg. Neg. Reg.	Proteins	Any	30
SITE	Binding Phosphorylation	Entity	Site	20
LOCATION	Localization	Entity	Localization	30

Results (official evaluation)



tasks / results		recall	precision	f-measure
Task 1	(f2m1)	28.63	20.88	24.15
	(f2m6)	27.18	23.92	25.45
Task 2	(f2m1)	25.02	18.32	21.15
	(f2m6)	24.49	21.63	22.97

Results / Events	(f2m1)			(f2m6)		
	p	r	fm	p	r	fm
Protein catabolism	78.6	55.0	64.7	71.4	55.6	65.5
Gene Expression	51.8	55.1	53.4	52.6	61.2	56.6
Phosphorylation	49.6	56.1	52.7	46.0	55.2	50.2
Transcription	48.9	19.8	28.1	38.7	29.6	33.5
Localization	28.2	42.9	34.0	23.3	48.9	33.3
Binding	19.5	12.1	14.9	22.4	14.4	17.5
Positive regulation	10.0	6.6	7.9	10.2	8.0	9.0
Negative regulation	9.8	7.9	8.8	7.9	7.7	7.8
Regulation	8.6	4.5	5.9	7.5	5.3	6.3

Results (alternative evaluation)

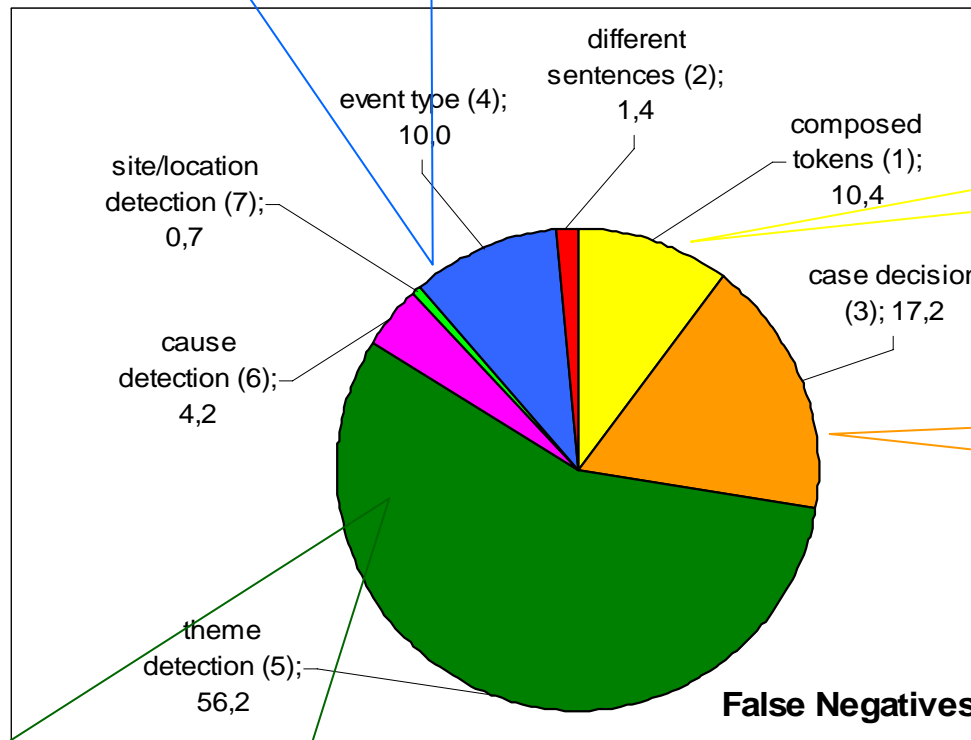


Events	(f2m1)			(f2m6)		
	p	r	fm	p	r	fm
Phosphorylation	75.0	94.7	83.7	79.1	89.5	84.0
Protein catabolism	70.8	89.5	79.1	69.6	84.2	76.2
Gene Expression	46.5	83.4	59.7	50.8	80.2	62.2
Localization	47.7	79.5	59.6	49.1	66.7	56.5
Transcription	22.7	75.9	34.9	36.4	74.6	48.9
Binding	29.7	71.1	41.9	29.7	64.4	40.7
Positive Regulation	24.3	63.7	35.2	26.5	59.1	36.6
Negative Regulation	26.4	56.5	36.0	25.3	43.5	32.0
Regulation	20.8	65.9	31.7	22.1	52.5	31.1
Entities (Site, Location, etc)	12.5	55.3	20.4	16.8	50.0	25.1
TOTAL	27.5	69.2	39.4	30.9	62.9	41.4

Mistakes (false negatives)



Overexpression [475-489], doc 10229231
Gene Expression and Positive Regulation



up-regulation [221-233],
doc 10411003

dysregulation [727-740],
doc 10229231

associated [278-288], doc 10196286
Themes: tumor necrosis factor receptor-associated factor (TRAF) 1 [294-351],
2 [353-354], 3 [359-360]

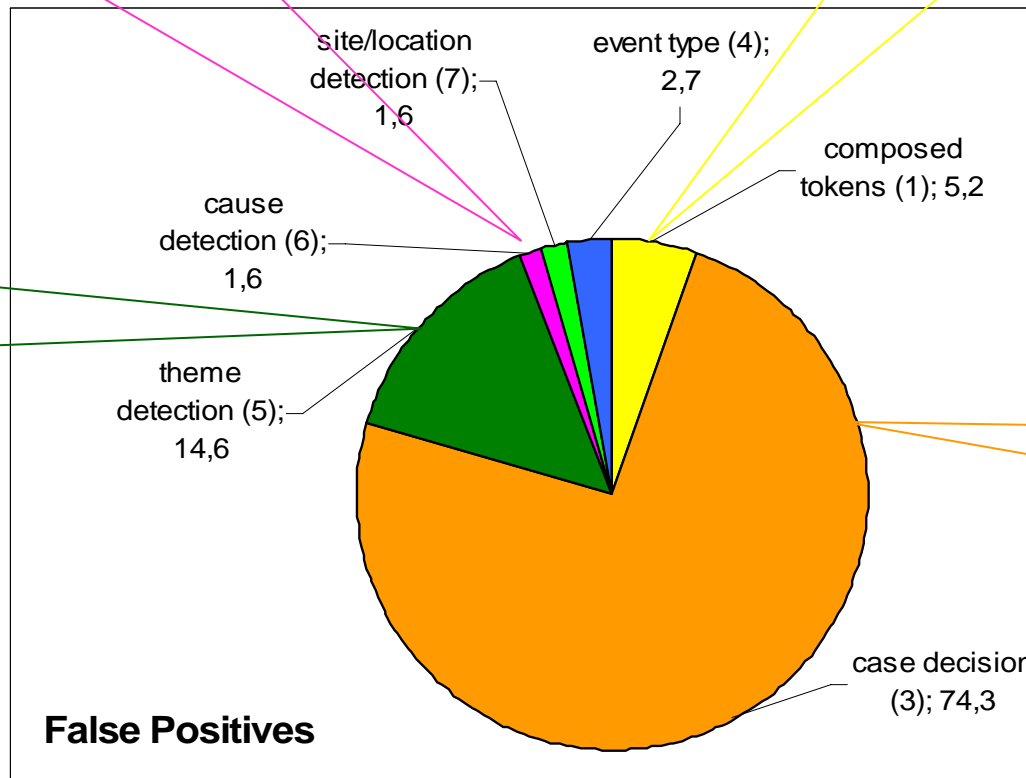
Mistakes (false positives)



decreases [1230-1239], Doc 10092805,
Theme 4E-BP1 [1240-1246],
Cause 4E-BP2 [1315-1321]

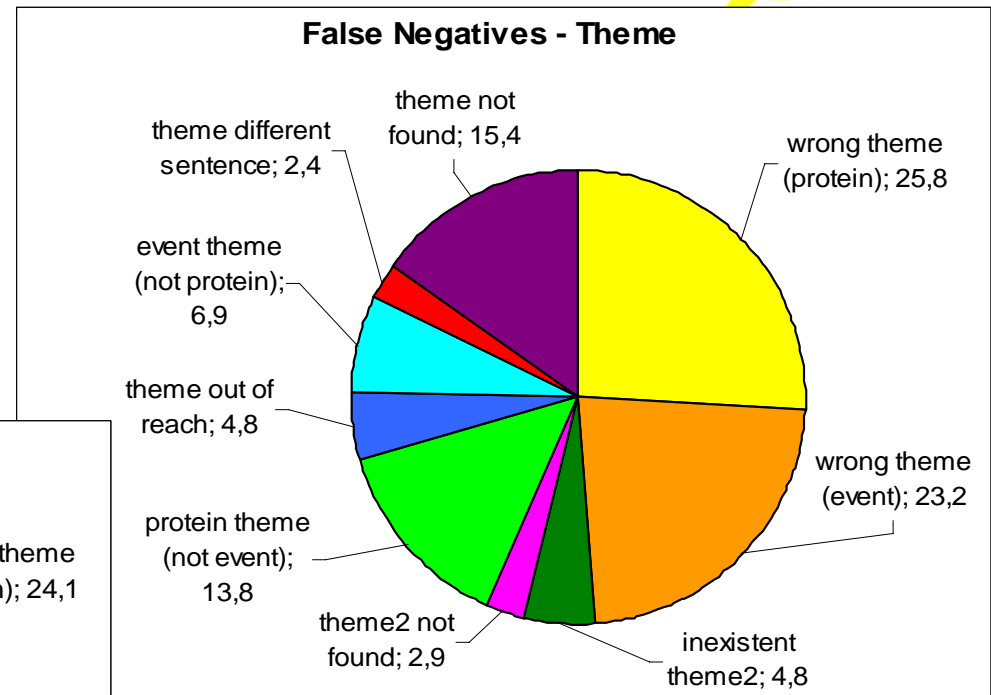
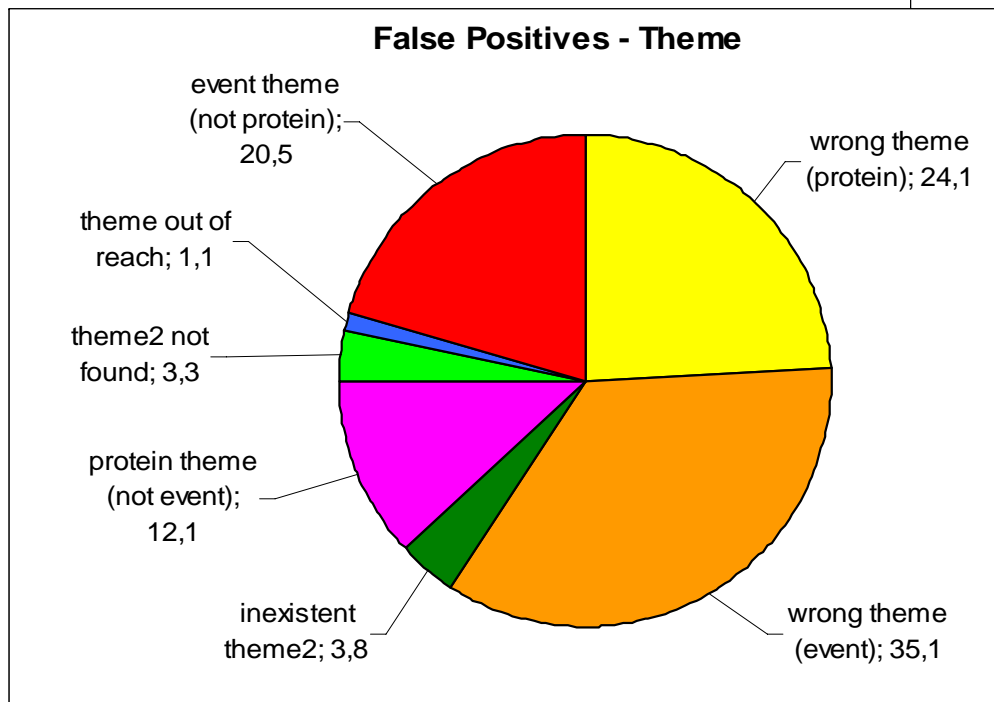
reduced levels [559-573],
doc 10411003

bound"[865-870],
doc 10089566,
Themes:
IL-10"[915-920],
gp41"[948-952]



transcript [392-402],
doc 10229231

Analysis of mistakes





Thank you.

Questions?

Mariana Lara Neves
mlara@cnb.csic.es
<http://moara.dacya.ucm.es/>