

# **The Usage of NooJ Resources for Detection of Certain Classes of Events**

---

**Sandra Gucul, Vanja Radulović,  
Cvetana Krstev  
Faculty of Philology  
University of Belgrade**

---

**INTRODUCTION**

---

An **information retrieval** task – our main **goal** is to retrieve some typical information from newspaper text using electronic dictionaries and finite state automata only.

Our **aim** is to retrieve as many occurrences that satisfy a query as possible and to exclude from the results as many occurrences as possible that do not satisfy a query.

# E-DICTIONARIES

Part-of-Speech (PoS) code

lemma, Cxxx[+SynSem] – various syntactic and semantic markers

code that determines its inflectional paradigm

# SEMANTIC MARKERS

semantic marker	value	
+First	first name	//Serbian/ English//
+Last	last name	//Serbian/ English//
+NProp	proper name	//Serbian/ English//
+Hum	human being	//Serbian/ English//
+Nick	nicknames	//Serbian/ English//
+SR	inhabitants of Serbia and Montenegro	//Serbian//
+EN	inhabitants of England	//English//
+Val	name as originally written	//English//
+Norm	correct transcription of the name	//English//

---

**SYSTEM OF GRAPHS**

---

- *more precise* formulation of the query
- *modulary design* –different composition of basic block enables various results

---

**MAIN BLOCKS OF OUR SYSTEM**

---

- *Persons*, that enable recognition of various references to persons
- *Dates*, that enable recognition of various references to time and date
- *Actions*, that detect actions undertaken by one or more persons at appropriate moment of time.

---

**PERSONS**

---

*Persons* – recognition of various references to person:

- full-name vs. surname
  - Serbian vs. English
  - male vs. female
  - name with or without the specified role
- etc.

---

**SIMPLE METHODS for personal names retrieval**

---

For retrieving **masculine** personal name:

$(\langle \mathbf{N+First:m} \rangle \langle \mathbf{N+Last:s} \rangle) + (\langle \mathbf{N+Last:s1} \rangle \langle \mathbf{N+First:m} \rangle)$

---

For retrieving **feminine** personal name:

$(\langle \mathbf{N+First:f} \rangle \langle \mathbf{N+Last:s1} \rangle) + (\langle \mathbf{N+Last:s1} \rangle \langle \mathbf{N+First:f} \rangle)$

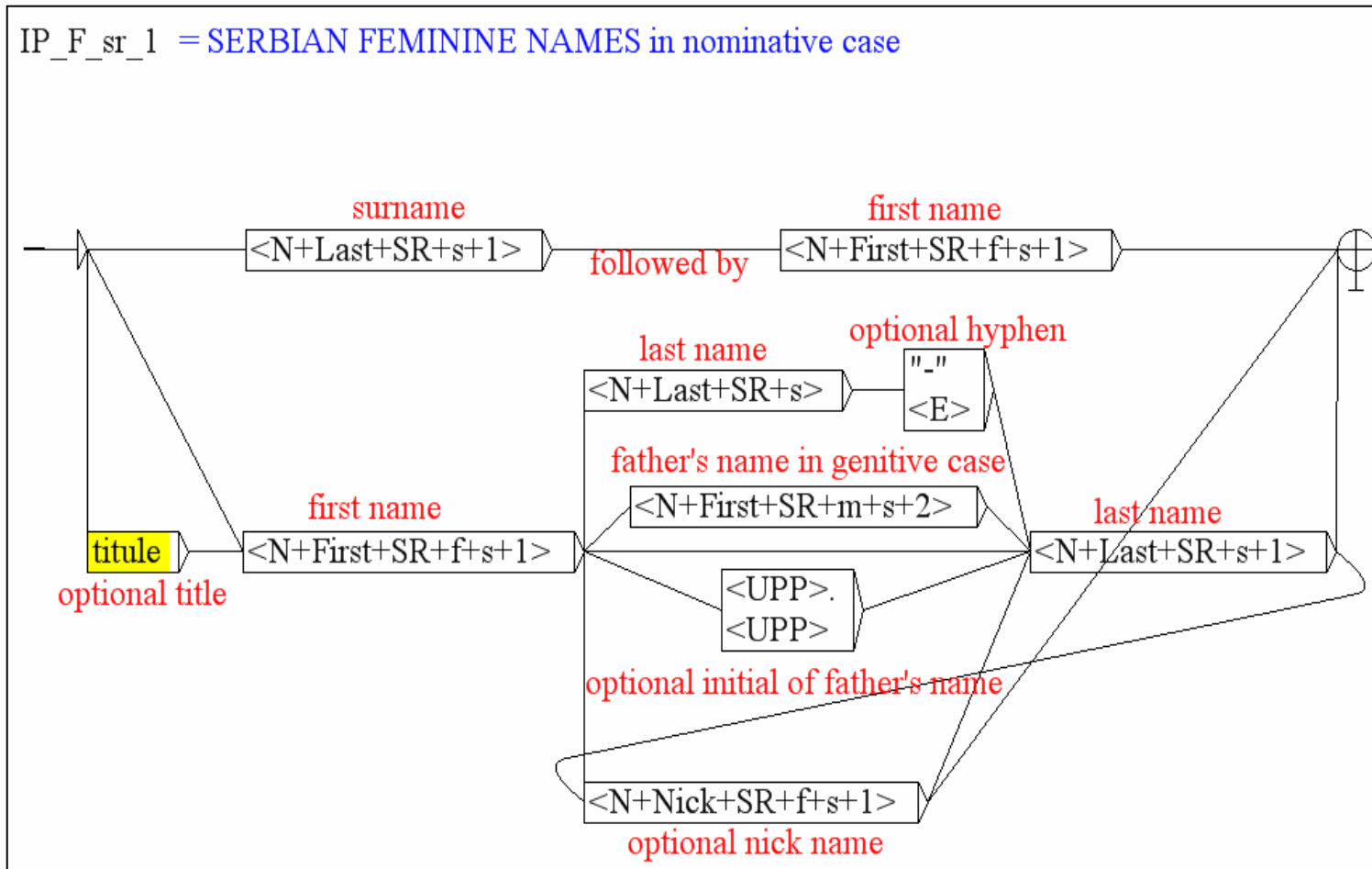
## FALSE occurrences retrieved by the query <N+First>

\*considering **homography** of personal names and **ambiguity** of many form\*

a) i crkava, potpuno kao kod nas.	<b>Andeli</b>	, nekada ljudi, ispisuju svoje misli na
b) košulja, cilindar, crn iberciger.	<b>Ide</b>	on tako i tetura se, i ja naletim na njega,
c) kao što znam ulice u Kadiksu.	<b>Divna</b>	stvar, to njihovo namesništvo! To
d) da osnuju školu za odrasle.	<b>Bila</b>	bi im potrebna jedna konferencija da

# ADVANCED methods for personal names retrieval

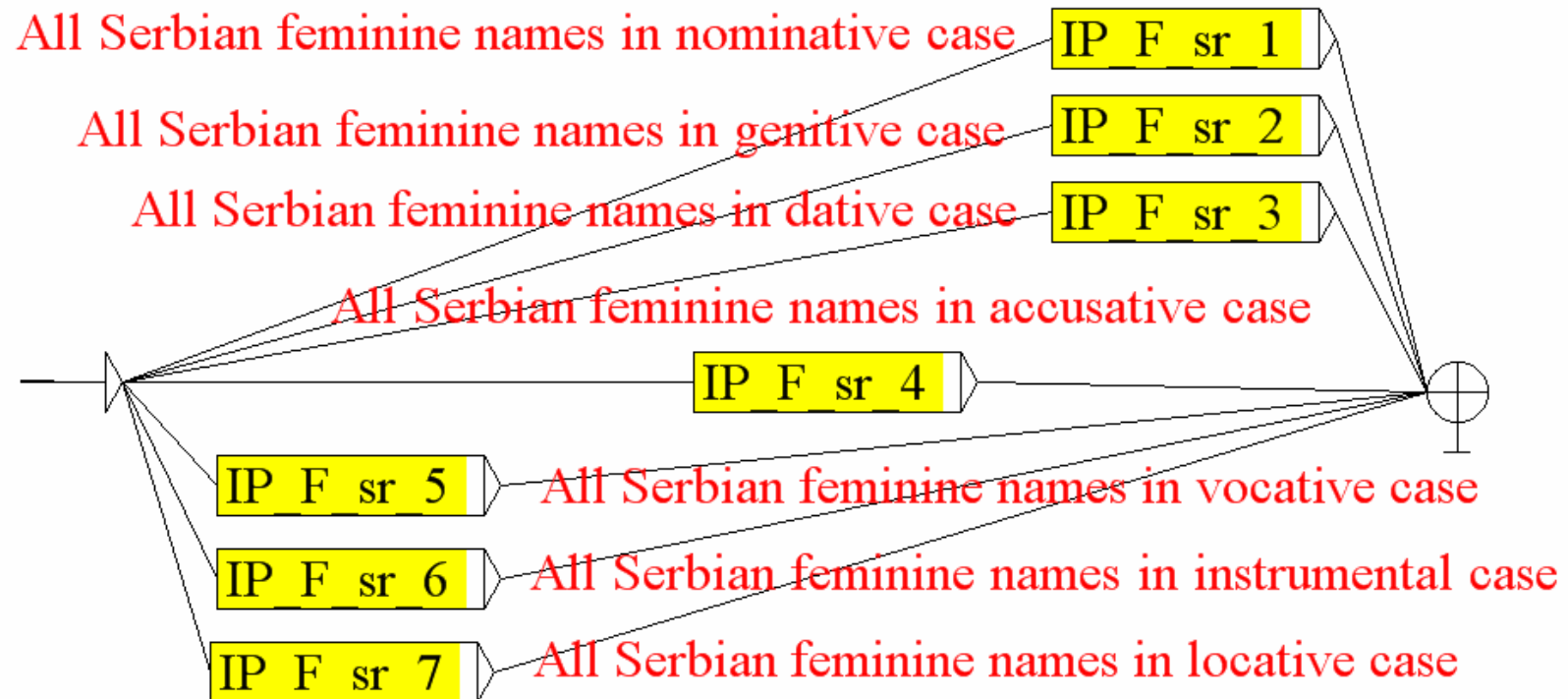
## basic subgraph – IP\_F\_sr\_1



## ADVANCED methods for personal names retrieval

$$IP\_F\_sr\_1 + IP\_F\_sr\_2 + \dots + IP\_F\_sr\_7 = Im\_Prez\_F\_sr$$

$Im\_Prez\_F\_sr =$  ALL SERBIAN FEMININE NAMES



## ADVANCED methods for personal names retrieval

$\text{Im\_Prez\_F\_sr} + \text{Im\_Prez\_M\_sr} = \text{Im\_Prez\_FM\_sr}$

$\text{Im\_Prez\_FM\_sr} = \text{ALL SERBIAN NAMES in all cases}$

All Serbian feminine names in all cases

Im Prez F sr

Im Prez M sr

All Serbian masculine names in all cases

## ADVANCED methods for personal names retrieval

$\text{Im\_Prez\_FM\_sr} + \text{Im\_Prez\_FM\_en} = \text{Im\_Prez\_FM\_all}$

$\text{Im\_Prez\_FM\_all} = \text{ALL FULL NAMES}$

All Serbian names in all cases

Im Prez FM sr

Im Prez FM en

All English names in all cases

**RESULT**

istar poljoprivrede  
 ministar finansija  
 m, po rečima prof.  
 ministra finansija  
 29. aprila guverner  
 no," kaže za B92  
 g košarkaškog asa  
 Narodne skupštine,  
 aj glavnog tužioca  
 esetih, u kojima je  
 objavlivanje pisma  
 arol ovog Društva,  
 eno značenje mu je  
 nik direktora EPS-a  
 eba imati u vidu da  
 torija finansija  
 xe brodova. Beranac  
 italačkoj publici.  
 xenjima?!", pita se  
 dnim pravom", kaže

Ivana Dulić-Marković  
 Mlađan Dinkić  
 dr Radomira Simića  
 Igor Lukšić  
 Radovan Jelašić  
 Milica Bisić  
 Vlade Divca  
 Boris Tadić  
 Karle Del Ponte  
 Džordž Lukas  
 Slobodana T. Jovanovića  
 dr Draško Karađinović  
 Josip Broz Tito.  
 Vladimir Đorđević  
 Bil Gejts  
 Vasilije J. Milić  
 Mihailo Milo Marković  
 Goran B. Milošević  
 Vida Petrović Škero  
 Vesna Rakić Vodinelić

saopštila je da 53 odsto  
 . "Tenderska procedur  
 sa Rudarsko-geološkog  
 izjavio je 29. aprila da  
 . Učestvujući u radu s  
 , zamenik ministra fina  
 o dokapitalizaciji 51 od  
 je u nedelju, 11. jula, z  
 koji saradnju Beograda  
 takođe tražio inspiracij  
 iz Beograda u oblast u  
 . "Bilo bi produženje k  
 Kako italijanske novin  
 . "Racionalizacioj glav  
 ni u najvećem bumu s  
 ,"Novac, kredit i banke  
 , nastavnik, kao i oblas  
 je kompleksnu oblast u  
 . Ona podsećah da je i  
 . Lice koje je listalo na

**SPECIFIC DEMANDS**

Subgraphs can be combined in **various ways**

Im\_Prez\_F\_all = ALL FEMININE NAMES IN ALL CASES

All Serbian feminine names in all cases

Im Prez F sr

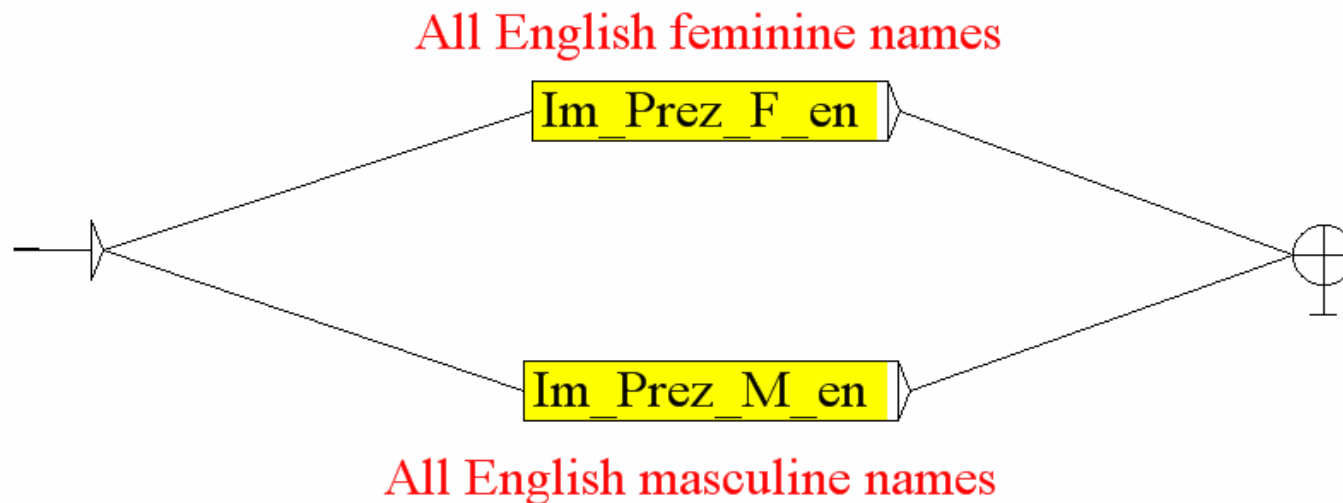
Im Prez F en

All English feminine names in all cases

**SPECIFIC DEMANDS**

Subgraphs can be combined in **various ways**

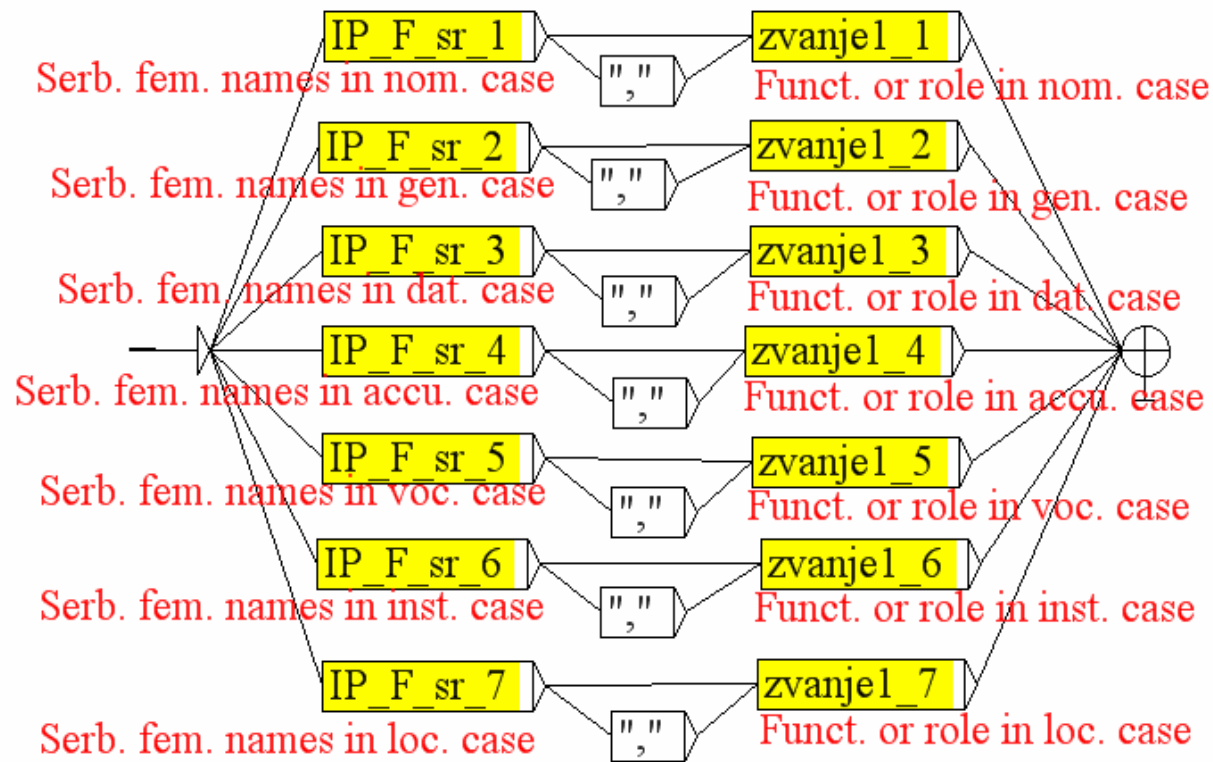
Im\_Prez\_FM\_en = ALL ENGLISH NAMES in all cases





# EXTRACTION of person's FUNCTION or ROLE

Im\_Prez\_F\_sr\_Z= FEMININE NAMES FOLLOWED BY THEIR FUNCTION



**RESULT**

ministarka poljoprivrede u Vladi Srbije Ivana Dulić-Marković
Vesna Džinić, direktor Agencije za osiguranje depozita, sanaciju, stečaj i likvidaciju banaka
zamenik srpskog ministra za međunarodne ekonomske odnose Duško Lopandić
Vojin Dimitrijević, direktor beogradskog Centra za ljudska prava i ekspert za međunarodno pravo
britanskog premijera Tonija Blera
Radoje Žugić, pomoćnik generalnog direktora Centralne banke Crne Gore za poslove platnog prometa
Ministar za međunarodne ekonomske odnose Srbije i Crne Gore Predrag Ivanović
Tončiju Staničiću, novom hrvatskom ambasadoru u Beogradu
gradonačelnik Čačka i lider Nove Srbije Velimir Ilić
Potpredsednik jugoslovenske vlade i savezni ministar za ekonomske odnose sa inostranstvom Miroljub Labus
američkog predsednika Džordža Buša
pomoćnik ministra za ekonomske odnose sa inostranstvom Ana Trbović
Dubravko Mikulić, direktor Sektora za poljoprivredu , prehrambenu industriju i šumarstvo Privredne komore Hrvatske
Rade Rakočević, predsednik Upravnog odbora poslovnog udruženja berzi i berzanskih posrednika
Ministar u Vladi Srbije za ekonomske odnose sa inostranstvom Predrag Bubalo
Boris Begović, potpredsednik Centra za liberalno demokratske studije i profesor ekonomije na beogradskom Pravnom fakultetu

---

**DATES**

---

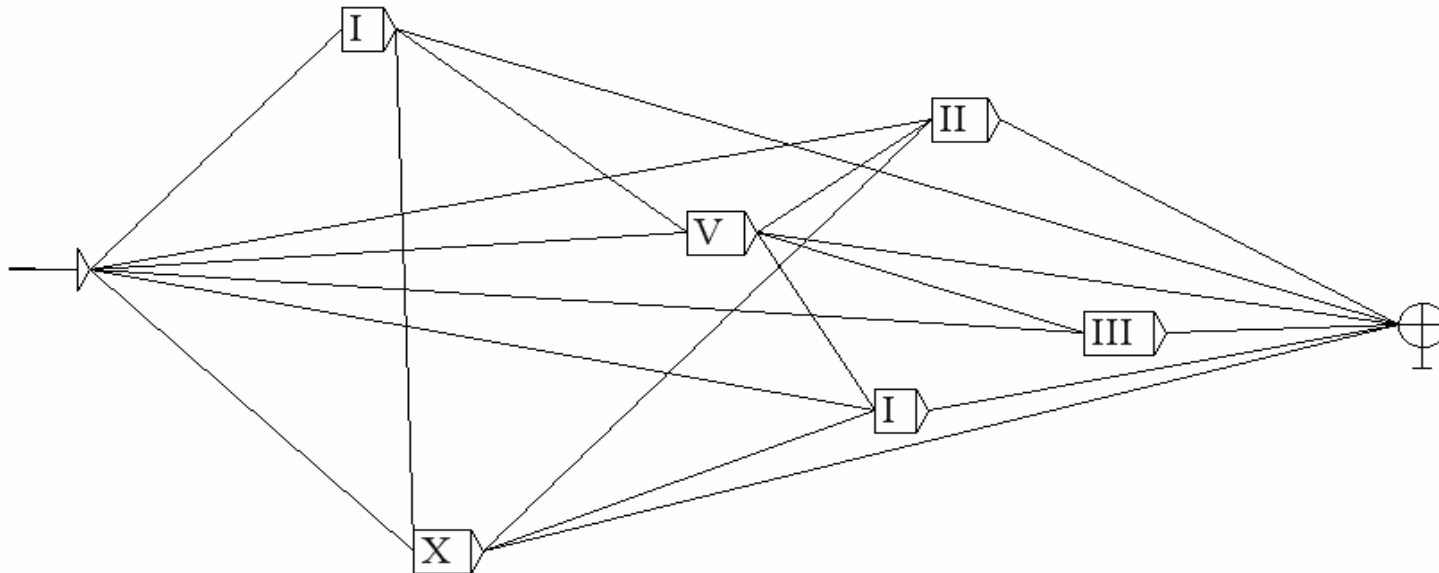
*Dates* – enables recognition of various references to time and date:

1. absolute vs. relative
2. complete vs. incomplete
3. specific demands

# ADVANCED METHODS for DATES retrieval

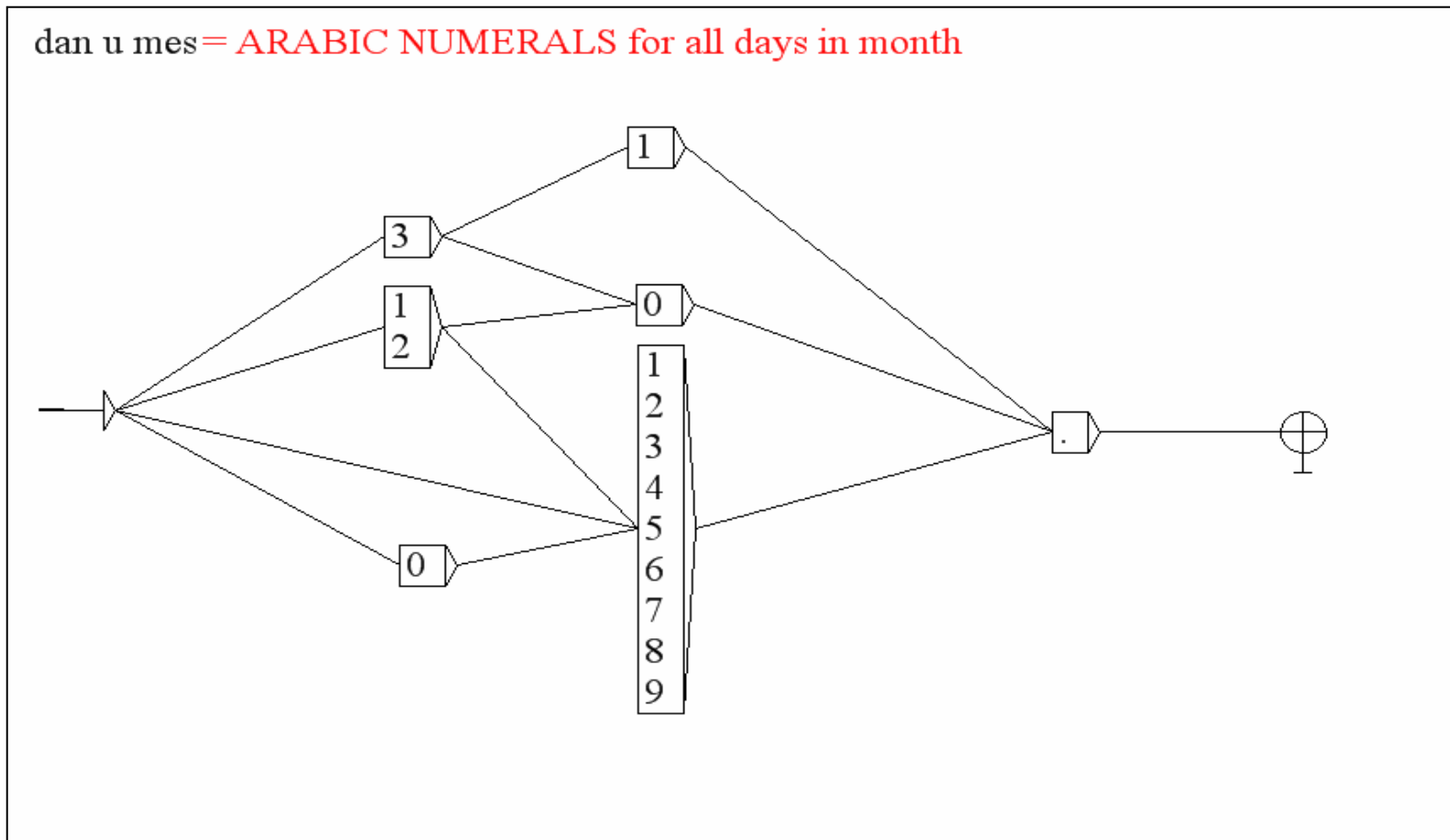
## *basic subgraph* – ROMAN NUMERALS

mes rimski = ROMAN NUMERALS for 12 months



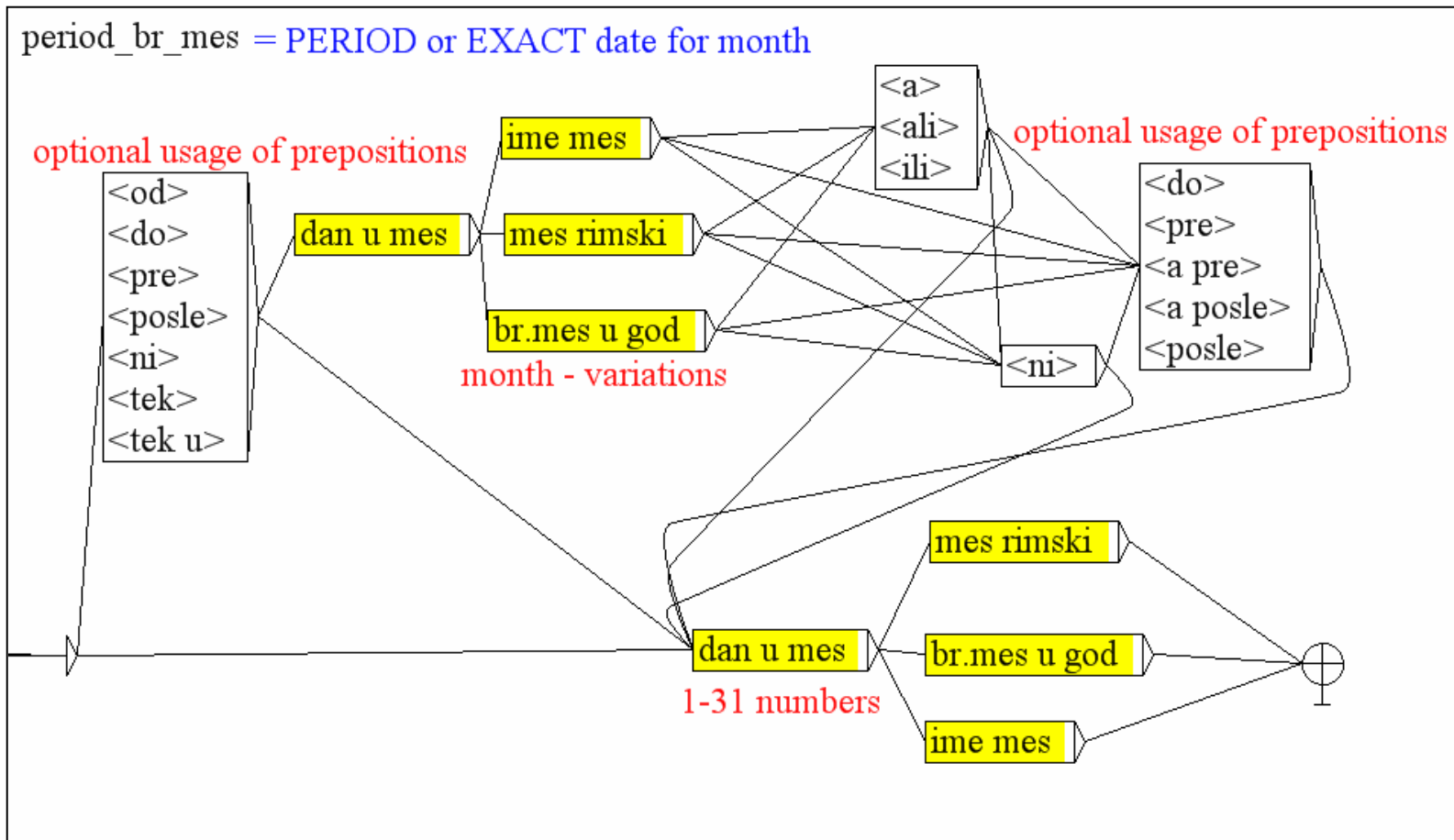
# ADVANCED METHODS for DATES retrieval

## *basic subgraph* – ARABIC NUMERALS



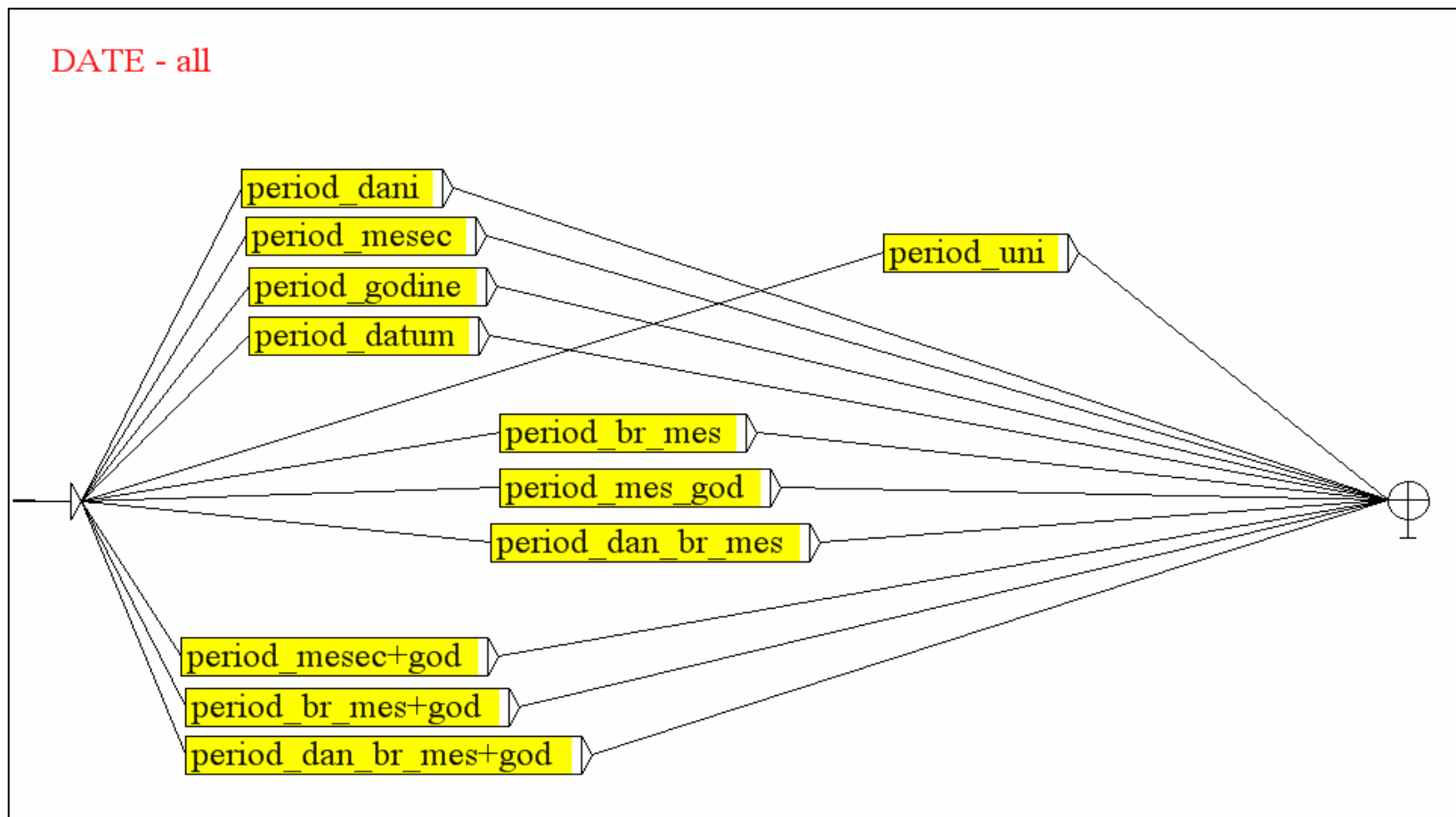
# ADVANCED METHODS for DATES retrieval

## PERIOD or EXACT date for month



# ADVANCED METHODS for DATES retrieval

## DATE ALL



**RESULT**

na U Makedoniji se  
e banke u Nikoziji,  
.000 evra i ostatak  
h proizvoda nastali  
ine biti isplaćene  
, dodavši da je od  
priraštaj štednje  
moraće, najkasnije  
nički dom Kongresa  
ansijskih eksperata  
predviđeno da HFP  
Hemofarm koncerna.  
dolazi u Hrvatsku.  
ti otvoren na jesen  
Đelić izjavio je  
na zarade i promet  
oizvoda u Bugarskoj  
a u evro-zoni opala  
štednje da bi ona  
2 odsto manje nego  
viru Programa vlade  
na u Evropsku uniju  
zetnike. "Do kraja

14. aprila  
22. februara 2001. godine,  
2007. godine.  
do 31. decembra 2002. godine.  
posle 20. februara  
prošle godine  
od januara do oktobra 2003. godine,  
do 1. marta,  
u decembru,  
u ponedeljak  
svakog meseca  
Tokom prošle godine  
Za utorak, 18. maja,  
sledeće godine  
u subotu, 13. marta,  
u 2004. godini.  
u prošloj godini  
tokom maja  
u narednim godinama  
u prethodnom mesecu  
za 2004. godinu.  
tek 2016. godine,  
prošle nedelje

održavaju predsednički  
bilo je 40,14 miliona d  
Određena sredstva Zaj  
Iz NIS-a podsećaju da  
. Za kilogram 'berleja' p  
počelo da se radi na reg  
prema MAP-u, bio je v  
da dostavi Narodnoj ba  
a sada ostaje da ga potp  
održati konferenciju za  
po 5 društava u kojima  
održano je 14 aukcija n  
planiran je nastavak raz  
. Slede Kragujevac, Niš  
da su vlasnici firmi "EF  
Ministar finansija i eko  
je bila 34,410 milijardi  
. Ministarstvo unutrašn  
postala dominantna kom  
. Podaci zvanične statist  
Beograd: Otvoren prvi f  
ili da bi prijem mogao  
Poreskoj upravi je stiglo

---

**ADVANCED METHODS for ACTION retrieval**

---

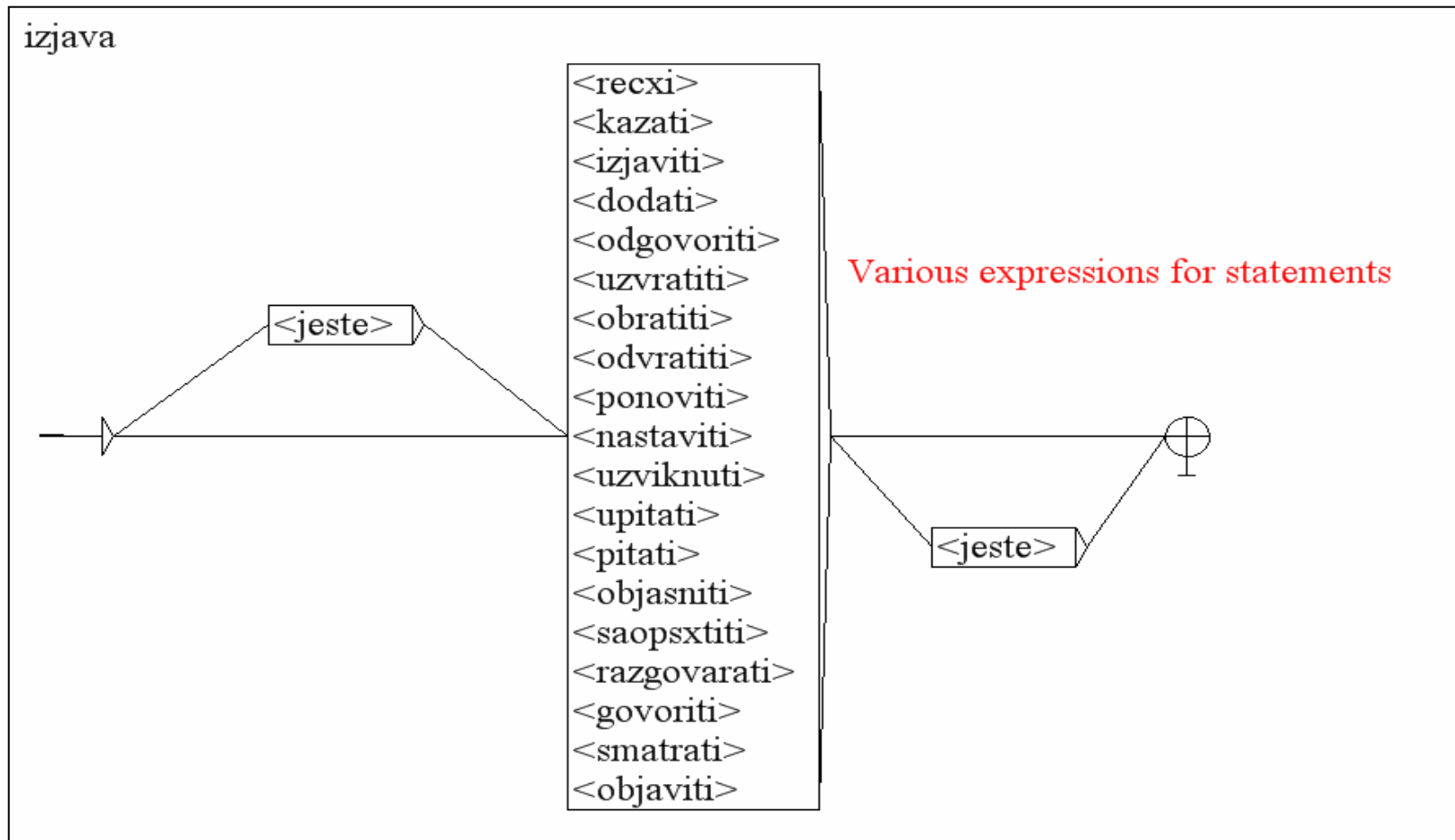
**Action** – detection of action undertaken by person(s) at appropriate moment of time:

1. statements

2. meetings

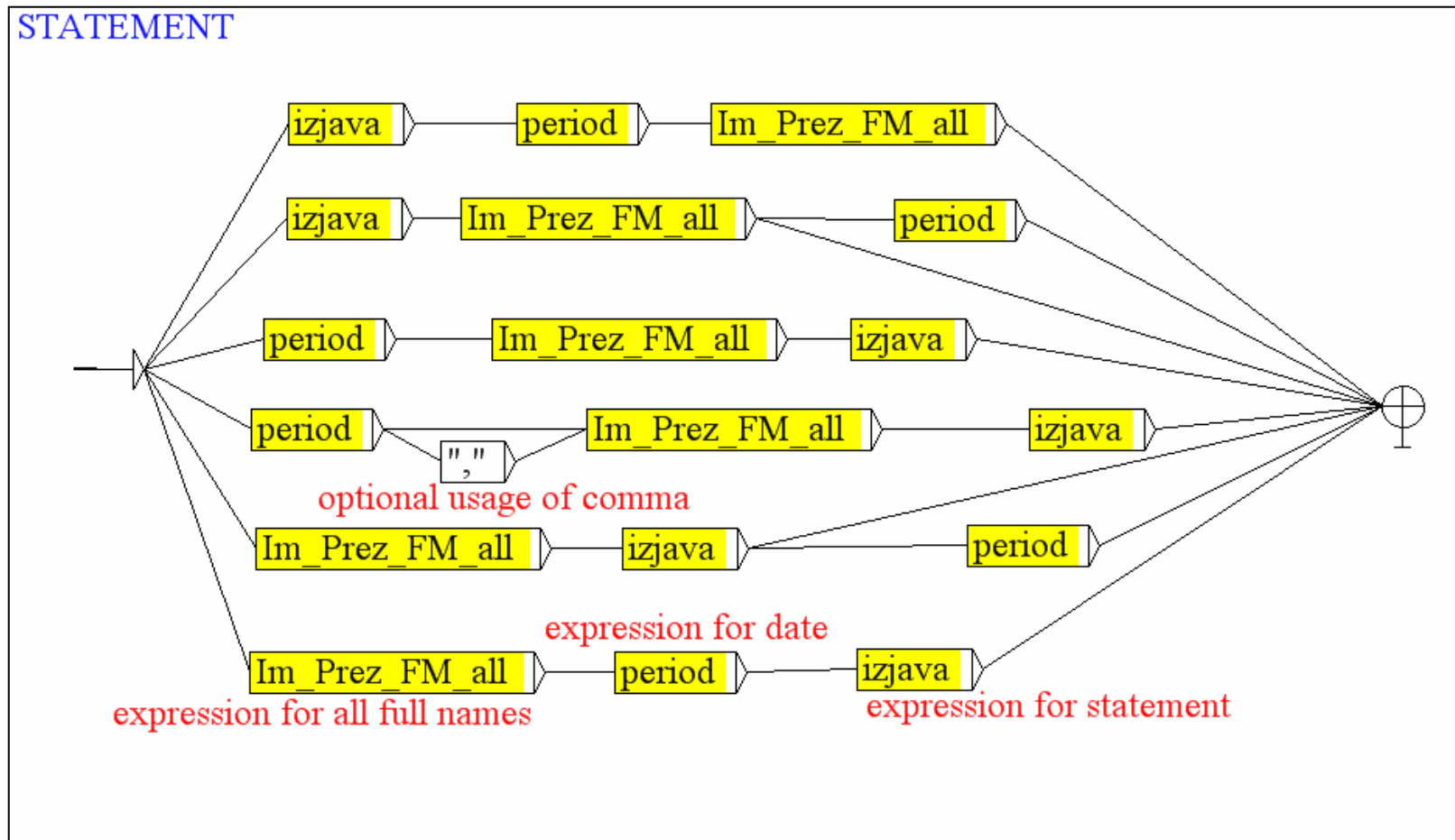
# STATEMENTS

## basic subgraph for statements extraction



# STATEMENTS

## main graph for statement extraction



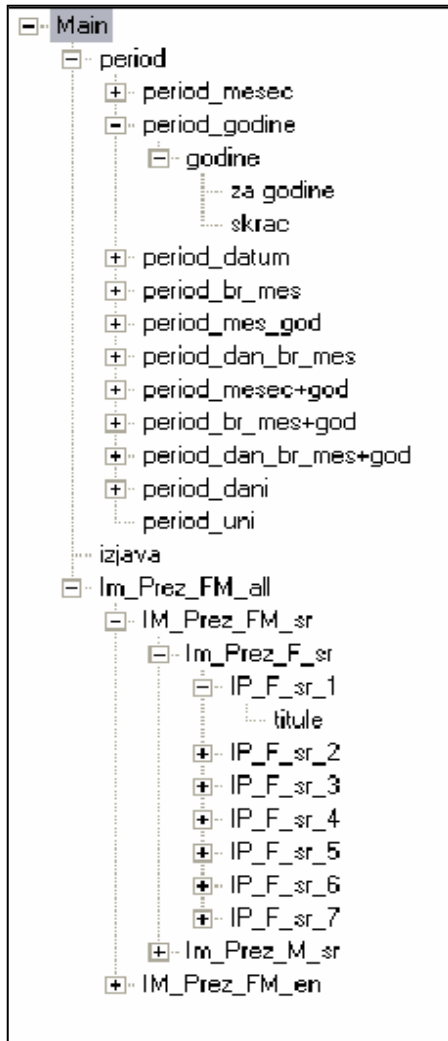
# RESULT

sistema "Svislajn" ministar finansija za još 15 odsto", socijalni položaj, crnogorski premijer rže pod kontrolom, Centralnog registra sto digitalizacije, bu protiv korupcije i ekonomije Srbije trgovine i turizma da za razvoj Srbije sednik vlade Srbije na evra. Ekonomista ministra finansija inansijski direktor sna i Hercegovina", na sudijskom poslu, yela u novom roku", Srbije za naš list

Rodoljub Drašković saopštio je danas Igor Lukšić saopštio je 8. juna saopštio je Mirosljub Labus smatra Vladimir Gligorov Milo Đukanović smatra rekao je Milorad Moračić Vida Uzelac je rekla izjavio je Draško Petrović Verica Barać u sredu je izjavila Božidar Đelić izjavio je u subotu, 13. marta, Slobodan Milosavljević izjavivši Olivera Božić izjavila je 27. juna Mirosljub Labus razgovarao je u subotu Ljubomir Madyar dodao je Koviljka Mihailović je kazala Dyon Konors kaže objasnio je Dragoljub Mićunović pita Vida Petrović Škero kaže Vida Petrović Škero govori Vladimir Kravčuk

da razmis da Crna Gora ne posle sastanka sa . Biće gore Danic da je postignuti ko , direktor Direkcij da je na spisku Pri , generalni direkto da očekuje da će S da su u Nišu da nije da je Fond u prv sa predstavnicim da je učešće subve da je ostvaren cilj da u kompaniji oč , koji predvodi d , u situaciji primen . Beda "Sudovi , direktor Merkato

# SYSTEM OF GRAPHS



Structure of the main graph  
for statement retrieval

---

**FUTURE WORK**

---

- expanding extraction of statements with persons function or role
- precise extraction of meetings that will include:
  - two or more persons
  - persons function or role
  - information of date of the meeting
  - information of place of the meeting
- quantification of obtained results.



THANK YOU