Computational Frameworks for Semantic Analysis and Wikification

Dan Roth
Department of Computer Science
University of Illinois at Urbana-Champaign

With thanks to:
Collaborators: Ming-Wei Chang, Xiao Cheng, Lev Ratinov, Vivek Srikumar, Many others
Funding: NSF; DHS; NIH; DARPA; IARPA; ARL
DASH Optimization (Xpress-MP); GUROBI Optimization
Please...
Learning and Inference

- Global decisions in which several local decisions play a role but there are mutual dependencies on their outcome.
  - In current NLP we often think about simpler structured problems: Parsing, Information Extraction, SRL, etc.
  - As we move up the problem hierarchy (Textual Entailment, QA,....) not all component models will be learned simultaneously
  - We need to think about (learned) models for different sub-problems
  - Knowledge relating sub-problems (constraints) becomes more essential and may appear only at evaluation time

- Goal: Reasoning that incorporates models’ information, along with prior knowledge, in making coherent decisions
  - Decisions that respect the local models as well as domain & context specific knowledge/constraints.
Relational Inference

...ousted long time Yugoslav President Slobodan Milošević in October. Mr. Milošević's Socialist Party

- What does Socialist Party refer to?

- There is a need to “look up” some information...
  - What and how to look up is determined by understanding local relations
  - These relations need to be coupled with relevant statistical models to support a decision

- How can we formulate this level of reasoning?
Constrained Conditional Models

\[
\arg\max_y \lambda \cdot F(x, y) - \sum_{i=1}^{K} \rho_i d(y, 1_C_i(x))
\]

**How to solve?**

This is an Integer Linear Program

Solving using ILP packages gives an exact solution.

Cutting Planes, Dual Decomposition & other search techniques are possible

**Inference** = best explanation (assignment) given models output & knowledge.

Penalty for violating the constraint.

(Soft) constraints component

Features, classifiers; log-linear models (HMM, CRF) or a combination

How far \( y \) is from a “legal” assignment

Weight Vector for “local” models

**How to train?**

**Training** is learning the objective function

Decouple? Decompose?

How to exploit the structure to minimize supervision?

Not Today

**Features, classifiers; log-linear models (HMM, CRF) or a combination**
Outline

- Integer Linear Programming Formulations for Natural Language Processing

- Example 1: Extended Semantic Role Labeling
  - Relaxing the pipeline
  - Dealing with lack of joint annotation: combining structured models

- Example 2: Wikification
  - Knowledge Acquisition by Grounding
  - Relational Inference for Wikification
  - Applications
Examples: CCM Formulations

$$\operatorname*{argmax}_y \lambda \cdot F(x, y) - \sum_{i=1}^{K} \rho_i d(y, 1_{C_i(x)})$$

CCMs can be viewed as a general interface to easily combine declarative domain knowledge with data driven statistical models

Formulate NLP Problems as ILP problems (inference may be done otherwise)

1. Sequence tagging \(\) (HMM/CRF + Global constraints)
2. Sentence Compression \(\) (Language Model + Global Constraints)
3. SRL \(\) (Independent Models + Global Constraints)

Constrained Conditional Models Allow:

- Learning a simple model \(\) (or multiple; or pipelines)
- Make decisions with a more complex model
- Accomplished by directly incorporating constraints to bias/re-rank global decisions composed of simpler models’ decisions
- More sophisticated algorithmic approaches exist to bias the output \(\) [CoDL: Cheng et. al 07,12; PR: Ganchev et. al. 10; DecL, UEM: Samdani et. al 12]
Semantic Role Labeling

I left my pearls to my daughter in my will.

\[ I \]_{A0} \ left \ [my \ pearls]_{A1} \ [to \ my \ daughter]_{A2} \ [in \ my \ will]_{AM-LOC}.

- \textbf{A0} Leaver
- \textbf{A1} Things left
- \textbf{A2} Benefactor
- \textbf{AM-LOC} Location

I left my pearls to my daughter in my will.
Algorithmic Approach

**Identify** argument candidates
- Pruning [Xue & Palmer, EMNLP’04]
- Argument Identifier
  - Binary classification

**Classify** argument candidates
- Argument Classifier
  - Multi-class classification

**Inference**

\[
\arg\max \sum_{a,t} y_{a,t} c_{a,t} = \sum_{a,t} 1_{a=t} c_{a=t}
\]

Subject to:
- One label per argument: \(\sum_t y_{a,t} = 1\)
- No overlapping or embedding
- Relations between verbs and arguments,....

No duplicate argument classes.

Learning Based Java: allows a developer to encode constraints in First Order Logic; these are compiled into linear inequalities automatically.

\[
\forall y \in \mathcal{Y}_R, \sum_{i=0}^{n-1} 1\{y_i = \text{"R-Ax"}\} \leq \sum_{i=0}^{n-1} 1\{y_i = \text{"Ax"}\}
\]

\[
\forall j, y \in \mathcal{Y}_C, 1\{y_j = \text{"C-Ax"}\} \leq \sum_{i=0}^{j} 1\{y_i = \text{"Ax"}\}
\]

Use the **pipeline architecture’s simplicity** while **maintaining uncertainty**: keep probability distributions over decisions & use global inference at decision time.

I left my nice pearls to her

I left my nice pearls to her
Improved sentence level analysis; dealing with more phenomena

BEIRUT, Lebanon — Lebanon’s main opposition group called for widespread protests on Sunday in the wake of a powerful bomb attack for which it blamed Syria, posing a challenge to a shaky coalition government that is led by pro-Syrian factions and intensifying fears that Syria’s civil war is spilling over into this country.

[Beirut] is in [Lebanon].
[Lebanon] has a main opposition group.
[Lebanon’s main opposition group] called for [widespread protests] [on Sunday].
There was [a powerful bomb attack].
[Lebanon’s main opposition group] blamed [Syria].
[Pro-Syrian factions] lead [a shaky coalition government]
[Syria] has a [civil war].
[Someone] fears that [Syria’s civil war is spilling over into this country].

Sentence level analysis may be influenced by other sentences
Verb SRL is not Sufficient

- John, a fast-rising politician, slept on the train to Chicago.

**Verb Predicate:** sleep

- **Sleeper:** John, a fast-rising politician
- **Location:** on the train to Chicago

**Who was John?**

- **Relation:** Apposition (comma)
- John, a fast-rising politician

**What was John’s destination?**

- **Relation:** Destination (preposition)
- train to Chicago
Example: preposition relations [Srikumar & Roth, 11,13]

Queen of England

City of Chicago
Predicates expressed by prepositions

- live at Conway House
- stopped at 9 PM
- drive at 50 mph
- look at the watch
- cooler in the evening
- the camp on the island
- arrive on the 9th

Index of definition on Oxford English Dictionary

- Location
- Temporal
- Numeric
- ObjectOfVerb

Ambiguity & Variability
Preposition relations [Transactions of ACL, ‘13]

- An inventory of 32 relations expressed by prepositions
  - Prepositions are assigned labels that act as predicate in a predicate-argument representation
  - Semantically related senses of prepositions merged
  - Substantial inter-annotator agreement

- A new resource: Word sense disambiguation data, re-labeled
  - SemEval 2007 shared task [Litkowski 2007]
    - ~16K training and 8K test instances; 34 prepositions
  - Small portion of the Penn Treebank [Dalhmeier, et al 2009]
    - only 7 prepositions, 22 labels

His first patient died of pneumonia. Another, who arrived from NY yesterday suffered from flu. Most others already recovered from flu.
Computational Questions

1. How do we predict the preposition relations?
   - Capturing the interplay with verb SRL (and eventually, others)
   - Very small jointly labeled corpus, cannot train a global model!

2. What about the arguments?
   - Annotation only gives us the predicate
   - How do we train an argument labeler?
   - Exploiting types as latent variables
Coherence of predictions

Predicate arguments from different triggers should be consistent

The bus was heading for Nairobi in Kenya.

Joint constraints
linking the two tasks.

Destination ⇔ A1

Predicate: head.02
A0 (mover): The bus
A1 (destination): for Nairobi in Kenya
Joint inference (CCMs)

Inference = best explanation (assignment) given models output & knowledge.

Variable $y^{a,t}$ is whether argument $a$ is assigned a label $t$.

$c^{a,t}$ is the corresponding model score.

Constraints:
- Each argument label
- Verb SRL constraints
- Only one label per preposition

+ Joint constraints between tasks; easy with ILP formulation

Joint Inference – no (or minimal) joint learning
Computational Questions

1. How do we predict the preposition relations?
   - Capturing the interplay with verb SRL (and eventually, others)
   - Very small jointly labeled corpus, cannot train a global model!

2. What about the arguments?
   - Annotation only gives us the predicate
   - How do we train an argument labeler?
   - Exploiting types as latent variables
Relations depend on argument types

- Our primary goal is to model preposition relations and their arguments.
- But the relation prediction strongly depends also on the semantic type of the arguments.

Poor care led to her death from pneumonia.

cause(death, pneumonia)

cause(death, flu)

Poor care led to her death from the flu.

The ability to generalize to unseen words of the same “type” would help argument & relation prediction.

Types are an abstraction that capture common properties of groups of entities.
Poor care led to her death from flu.

Prediction $y$:

Type $h(y)$

Relation $r(y)$

Governor type

Object type

cause

disease

death

flu

experience

Requires Latent Inference, incorporating constraints between hidden and supervised parts of the structure $y = (r, h)$, formulated as an ILP.

Types are represented as hidden variables that correspond to wordnet layer and distributional clusters.
Performance: The More (hidden) Components the Better

- Model size: 2.21 non-zero weights
- Model size: 5.41 non-zero weights

- Learned to predict both predicates and arguments
- Using types helps. Joint inference with word sense helps more.
- More components constrain inference results and improve performance

- Initialization
- + Latent
Extended SRL

<table>
<thead>
<tr>
<th></th>
<th>Preposition</th>
<th>Preposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>The bus was</td>
<td></td>
<td></td>
</tr>
<tr>
<td>heading to</td>
<td>V: head</td>
<td>Governor</td>
</tr>
<tr>
<td>Nairobi in</td>
<td>Destination</td>
<td>Object</td>
</tr>
<tr>
<td>Kenya</td>
<td>Location</td>
<td>Object</td>
</tr>
</tbody>
</table>

Joint inference over phenomena specific models to enforce consistency

- More to do with other relations, discourse phenomena,...
Constrained Conditional Models—ILP Formulations

- Have been shown useful in the context of many NLP problems

- [Roth&Yih, 04,07: Entities and Relations; Punyakanok et. al: SRL ...]
  - Summarization; Co-reference; Information & Relation Extraction; Event Identifications; Transliteration; Textual Entailment; Knowledge Acquisition; Sentiments; Temporal Reasoning, Dependency Parsing,...

- Some theoretical work on training paradigms [Punyakanok et. al., 05 more; Constraints Driven Learning, PR, Constrained EM...]

- Some work on Inference, mostly approximations, bringing back ideas on Lagrangian relaxation, etc.

- Good summary and description of training paradigms: [Chang, Ratinov & Roth, Machine Learning Journal 2012]

- Summary of work & a bibliography: http://L2R.cs.uiuc.edu/tutorials.html
Outline

- Integer Linear Programming Formulations for Natural Language Processing

Example 1: Extended Semantic Role Labeling
  - Relaxing the pipeline
  - Dealing with lack of joint annotation: combining structured models

Example 2: Wikification
  - Knowledge Acquisition by Grounding
  - Relational Inference for Wikification
  - Applications
Blumenthal (D) is a candidate for the U.S. Senate seat now held by Christopher Dodd (D), and he has held a commanding lead in the race since he entered it. But the Times report has the potential to fundamentally reshape the contest in the Nutmeg State.
Applications

- Knowledge Acquisition via Grounding
- Coreference Resolution
  - Learning-based multi-sieve co-reference resolution with knowledge (Ratinov et al. 2012)
- Information Extraction
  - Unsupervised relation discovery with sense disambiguation (Yao et al. 2012)
  - Automatic Event Extraction with Structured Preference Modeling (Lu and Roth, 2012)
- Text Classification
  - Gabrilovich and Markovitch, 2007; Chang et al., 2008
- Entity Linking
Challenges

Blumenthal (D) is a candidate for the U.S. Senate seat now held by Christopher Dodd (D), and he has held a commanding lead in the race since he entered it. But the Times report has the potential to fundamentally reshape the contest in the Nutmeg State.

- State-of-the-art systems (Ratinov et al. 2011) can achieve the above with local and global statistical features
  - Reaches bottleneck around 70%~ 85% F1 on non-wiki datasets
  - Check out our demo at: http://cogcomp.cs.illinois.edu/demos
  - What is missing?
Relational Inference

- Mubarak, the wife of deposed Egyptian President Hosni Mubarak,…
Relational Inference

Mubarak, the wife of deposed Egyptian President Hosni Mubarak,...

- What are we missing with Bag of Words (BOW) models?
  - Who is Mubarak?
- Textual relations provide another dimension of text understanding
- Can be used to constrain interactions between concepts
  - (Mubarak, wife, Hosni Mubarak)
- Has impact in several steps in the Wikification process:
  - From candidate selection to ranking and global decision
Relational Inference for Wikification

- **Mubarak**, the wife of deposed Egyptian President **Hosni Mubarak**, ...

Next we will briefly show:

- How to identify key textual relations for Wikification
- How to verify relations using external resource
- A global inference framework to incorporate relational knowledge

Relational inference yields significant improvements over state-of-the-art systems
Wikification

- Mention Detection
- Candidate Generation
- Candidate Ranking
- Determine NILs
Wikification Pipeline 1 - Mention Detection

...ousted long time Yugoslav President Slobodan Milošević in October. Mr. Milošević's Socialist Party...

- sub-NP (Noun Phrase) chunks [Illinois Chunker]
- NER [Illinois NER]
- Regular expressions
...ousted long time Yugoslav President Slobodan Milošević in October. Mr. Milošević's Socialist Party...
...ousted long time Yugoslav President Slobodan Milošević in October. Mr. Milošević's Socialist Party...
...ousted long time Yugoslav President Slobodan Milošević in October. Mr. Milošević's Socialist Party...

<table>
<thead>
<tr>
<th>k</th>
<th>e^k_3</th>
<th>s^k_3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Slobodan_Milošević</td>
<td>0.7</td>
</tr>
<tr>
<td>2</td>
<td>Milošević_(surname)</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>Boki_Milošević</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>Alexander_Milošević</td>
<td>0.05</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>k</th>
<th>e^k_4</th>
<th>s^k_4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Socialist_Party_(France)</td>
<td>0.23</td>
</tr>
<tr>
<td>2</td>
<td>Socialist_Party_(Portugal)</td>
<td>0.16</td>
</tr>
<tr>
<td>3</td>
<td>Socialist_Party_of_America</td>
<td>0.07</td>
</tr>
<tr>
<td>4</td>
<td>Socialist_Party_(Argentina)</td>
<td>0.06</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Local and global statistical features
...ousted long time **Yugoslav President** Slobodan Milošević in October. Mr. Milošević’s **Socialist Party**...

<table>
<thead>
<tr>
<th>k</th>
<th>e&lt;sub&gt;k&lt;/sub&gt;</th>
<th>s&lt;sub&gt;k&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Slobodan_Milošević</td>
<td>0.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>k</th>
<th>e&lt;sub&gt;k&lt;/sub&gt;</th>
<th>s&lt;sub&gt;k&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Socialist_Party_(France)</td>
<td>0.23</td>
</tr>
</tbody>
</table>

- Is the top candidate really what the text referred to?
  - If **NO**, no title is assigned to this mention.
**Formulation**

- **Goal:** Promote concepts that are coherent with textual relations

- Formulate as an Integer Linear Program (ILP):

\[
\Gamma_D = \arg \max \sum \sum s_i^k e_i^k + \sum \sum w_{ij}^{(k,l)} r_{ij}^{(k,l)}
\]

\[
s.t. \quad r_{ij}^{(k,l)} \in \{0, 1\} \quad \text{Integral constraints}
\]

\[
e_i^k \in \{0, 1\} \quad \text{Integral constraints}
\]

\[
\forall i \sum_k e_i^k = 1 \quad \text{Unique solution}
\]

\[
2r_{ij}^{(k,l)} \leq e_i^k + e_j^l \quad \text{Relation definition}
\]

- If no relation exists, collapses to the non-structured decision
...ousted long time **Yugoslav President** Slobodan Milošević in October. Mr. **Milošević’s Socialist Party**...

<table>
<thead>
<tr>
<th>k</th>
<th>Concept</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Slobodan_Milošević</td>
<td>0.7</td>
</tr>
<tr>
<td>2</td>
<td>Milošević_(surname)</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>Boki_Milošević</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>Alexander_Milošević</td>
<td>0.05</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>k</th>
<th>Concept</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Socialist_Party_(France)</td>
<td>0.23</td>
</tr>
<tr>
<td>2</td>
<td>Socialist_Party_(Portugal)</td>
<td>0.16</td>
</tr>
<tr>
<td>3</td>
<td>Socialist_Party_of_America</td>
<td>0.07</td>
</tr>
<tr>
<td>4</td>
<td>Socialist_Party_(Argentina)</td>
<td>0.06</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\Gamma_D = \arg \max_{\Gamma} \sum_i \sum_k s_i^k e_i^k + \sum_{i,j} \sum_{k,l} w_{ij}^{(k,l)} r_{ij}^{(k,l)}
\]

- \(e_i^k\): whether a concept is chosen
- \(s_i^k\): score of a concept
- \(r_{ij}^{(k,l)}\): whether a relation is present
- \(w_{ij}^{(k,l)}\): score of a relation
Overall Approach

Wikification

- Candidate Generation
- Candidate Ranking
- Determine NILs

Relation Analysis

- Relation Identification
- Relation Retrieval
- Relational Inference
1. Relation Identification

- **ACE style in-document coreference** [Chang et. al, EMNLP’13]
  - Extract named entity-only coreference relations with high precision

- **Syntactico-Semantic relations** [Chan & Roth ’10]

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premodifier</td>
<td>Iranian <strong>Ministry of Defense</strong></td>
</tr>
<tr>
<td>Possessive</td>
<td>NYC’s <strong>stock exchange</strong></td>
</tr>
<tr>
<td>Formulaic</td>
<td><strong>Chicago, Illinois</strong></td>
</tr>
<tr>
<td>Preposition</td>
<td><strong>President</strong> of the <strong>US</strong></td>
</tr>
</tbody>
</table>

- Easy to extract with high precision
- Aim for high recall, as false-positives will be verified and discarded
- These relations covers ~80% relation instances in ACE2004
1. Relation Identification

...ousted long time **Yugoslav President** Slobodan Milošević in October. Mr. **Milošević**’s **Socialist Party**...

<table>
<thead>
<tr>
<th>Argument 1</th>
<th>Relation Type</th>
<th>Argument 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yugoslav President</td>
<td>apposition</td>
<td>Slobodan Milošević</td>
</tr>
<tr>
<td>Slobodan Milošević</td>
<td>coreference</td>
<td>Milošević</td>
</tr>
<tr>
<td>Milošević</td>
<td>possessive</td>
<td>Socialist Party</td>
</tr>
</tbody>
</table>
2. Relation Retrieval for Candidate Generation

...ousted long time Yugoslav President Slobodan Milošević in October. Mr. Milošević's Socialist Party...

- Earlier approach
  - Collect known mappings from Wikipedia page titles, hyperlinks...
  - Limit to top-K candidates based on frequency of links (Ratinov et al. 2011)

- What concepts can “Socialist Party” refer to?
A Lot of Uninformative Mentions

Socialist Party (disambiguation)

From Wikipedia, the free encyclopedia

Socialist Party is the name of many different political parties and article.

Socialist Party may also refer to the wide variety of political parties. What follows is an incomplete alphabetical list of such parties:

Names used by several different parties [edi]

- Arab Socialist Ba'ath Party (disambiguation)
- Authentic Socialist Party (disambiguation)
- Democratic Socialist Party (disambiguation)
- Independent Socialist Party (disambiguation)
- National Socialist Party (disambiguation)
- New Socialist Party (disambiguation)
- Polish Socialist Party (disambiguation)
- Popular Socialist Party (disambiguation)
- Revolutionary Socialist Party (disambiguation)
- Socialist Action Party (disambiguation)
- Socialist Democratic Party (disambiguation)
- Socialist Equality Party (disambiguation)
- Socialist Labor Party (disambiguation)
- Socialist Labour Party (disambiguation)
- Socialist Left Party (disambiguation)
- Socialist People's Party (disambiguation)
- Socialist Republican Party (disambiguation)
- Socialist Unity (disambiguation)
2. Relation Retrieval for Candidate Generation

...ousted long time Yugoslav President Slobodan Milošević in October. Mr. Milošević's Socialist Party...

- What concepts can “Socialist Party” refer to?
  - Recall that the top ranked one is Socialist_Party_(France)

- More robust candidate generation
  - All identified relations are verified against a knowledge base (DBPedia)
  - Retrieve relation arguments matching “(Milošević,?,Socialist Party)” as our new candidates
2. Relation Retrieval for Candidate Generation

...ousted long time **Yugoslav President** Slobodan Milošević in October. Mr. **Milošević**'s **Socialist Party**...

<table>
<thead>
<tr>
<th>k</th>
<th>$e^k_3$</th>
<th>$s^k_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Slobodan_Milošević</td>
<td>0.7</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>k</th>
<th>$e^k_4$</th>
<th>$s^k_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Socialist_Party_(France)</td>
<td>0.23</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

- **Query Pruning**
  - Only 2 queries per pair necessary due to strong baseline.
  - 8 Relation: (Candidate, Related Segment); (Related Segment, Candidate)

$q_1=(**Socialist Party of France**, ?, *Milošević*)$
$q_2=(**Slobodan Milošević**, ?, *Socialist Party*)$
2. Relation Retrieval for Candidate Generation

<table>
<thead>
<tr>
<th>Argument 1</th>
<th>Relation Type</th>
<th>Argument 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milošević</td>
<td>possessive</td>
<td>Socialist Party</td>
</tr>
</tbody>
</table>

Milošević

Personal details

Born 20 August 1941
Požarevac, Yugoslavia
Died 11 March 2006 (aged 64)
The Hague, Netherlands
Nationality Serbian
Political party Socialist Party of Serbia (after 1990)
League of Communists of Yugoslavia (until 1990)
Spouse(s) Mirjana Marković
Children Marko and Marija
Alma mater University of Belgrade Faculty of Law
Religion Atheist
Signature

Socialist Party of Serbia
Социјалистичка Партија Србије
Socijalistička Partija Srbije

President Ivica Dačić
Founder Slobodan Milošević
Founded 17 July 1990
Preceded by League of Communists of Serbia
2. Relation Retrieval for Candidate Generation

...ousted long time **Yugoslav President** Slobodan Milošević in October. Mr. **Milošević's Socialist Party**...

<table>
<thead>
<tr>
<th>k</th>
<th>e₃^k</th>
<th>s₃^k</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Slobodan_Milošević</td>
<td>0.7</td>
</tr>
<tr>
<td>2</td>
<td>Milošević_(surname)</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>Boki_Milošević</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>Alexander_Milošević</td>
<td>0.05</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>k</th>
<th>e₄^k</th>
<th>s₄^k</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Socialist_Party_(France)</td>
<td>0.23</td>
</tr>
<tr>
<td>2</td>
<td>Socialist_Party_(Portugal)</td>
<td>0.16</td>
</tr>
<tr>
<td>3</td>
<td>Socialist_Party_of_America</td>
<td>0.07</td>
</tr>
<tr>
<td>4</td>
<td>Socialist_Party_(Argentina)</td>
<td>0.06</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Socialist_Party_of_Serbia</td>
<td>0.0</td>
</tr>
</tbody>
</table>
3. Relational Inference For Candidate Ranking

...ousted long time **Yugoslav President** Slobodan Milošević in October. Mr. **Milošević**'s **Socialist Party**...

<table>
<thead>
<tr>
<th>k</th>
<th>$e^k_3$</th>
<th>$s^k_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Slobodan_Milošević</td>
<td>0.7</td>
</tr>
<tr>
<td>2</td>
<td>Milošević_(surname)</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>Boki_Milošević</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>Alexander_Milošević</td>
<td>0.05</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>k</th>
<th>$e^k_4$</th>
<th>$s^k_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Socialist_Party_(France)</td>
<td>0.23</td>
</tr>
<tr>
<td>2</td>
<td>Socialist_Party_(Portugal)</td>
<td>0.16</td>
</tr>
<tr>
<td>3</td>
<td>Socialist_Party_of_America</td>
<td>0.07</td>
</tr>
<tr>
<td>4</td>
<td>Socialist_Party_(Argentina)</td>
<td>0.06</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Socialist_Party_of_Serbia</td>
<td>0.0</td>
</tr>
</tbody>
</table>

$$r_{34} \gamma(1,21) = 1$$

$$w_{34} \gamma(1,21) = ?$$

$$\Gamma_D = \arg \max_{\Gamma} \sum_{i} \sum_{k} s^k_i e^k_i + \sum_{i,j} \sum_{k,l} w^{(k,l)}_{ij} r^{(k,l)}_{ij}$$
3. Relational Inference For Candidate Ranking - Coreference

...ousted long time **Yugoslav President** Slobodan Milošević in October. Mr. Milošević's **Socialist Party**...

<table>
<thead>
<tr>
<th>k</th>
<th>(e^k_3)</th>
<th>(s^k_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Slobodan_Milošević</td>
<td>0.7</td>
</tr>
<tr>
<td>2</td>
<td>Milošević_(surname)</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>Boki_Milošević</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>Alexander_Milošević</td>
<td>0.05</td>
</tr>
</tbody>
</table>

\[ r \Downarrow 23 \Uparrow(1,1) \]

- Ranking is propagated via other relations to other candidates
4. Relation Inference for Determining Unknown Concepts

Dorothy Byrne, a state coordinator for the Florida Green Party,

How to capture the fact that:
- “Dorothy Byrne” does not refer to any concept in Wikipedia

Identify coreferent nominal mention relations
- Generate better features for NIL classifier and propagate via inference framework
Wikification Performance Result

F1 Performance on Wikification datasets

- ACE
- MSNBC
- AQUAINT
- Wikipedia

- Milne&Witten
- Ratinov&Roth
- Relational Inference
Evaluation – TAC KBP Entity Linking

- Run Relational Inference (RI) Wikifier “as-is”:
  - No retraining using TAC data

### TAC KBP 2011 Entity Linking Performance

![Bar chart showing TAC KBP 2011 Entity Linking Performance](chart)

- **System Names**
  - LCC
  - MS_MLI
  - RI
  - NUShine
  - CogComp
  - Median

- **Performance Metrics**
  - Micro Average
  - B³F1

*Median of top 14 systems*
Conclusion

- Presented **Constrained Conditional Models**
  - A powerful & modular learning and inference paradigm for high level tasks.
  - **Inference** for the best global explanation given (declarative) knowledge

- An ILP based computational framework that provides an interface to augment *statistically learned linear models* with *declarative constraints*
  - Incorporating knowledge and support decisions in expressive output spaces
  - Flexibility in Training & Inference [E.g., Amortized Inference, ACL’13, EMNLP’12]

- Exemplified the use of CCM in the context of layers of semantic inferences
  - Extended Semantic Role Labeling of Sentences
  - Wikification

Check out our tools, demos, LBJ and CCM tutorial