PEKING UNIVERSITY

Integrating Relation Constraints with Neural Relation Extractors

Yuan Ye, Yansong Feng*, Bingfeng Luo, Yuxuan Lai, Dongyan Zhao

Wangxuan Institute of Computer Technology, Peking University

2019/12/22

Relation Extraction

Relation Extraction (RE) aims to extract predefined relations between two marked entities in plain texts.

Barack Obama married Michelle Obama on October 3, 1992.

$\underbrace{\textit{Barack Obama}}_{\textit{Entity}_1}$ *married* $\underbrace{\textit{Michelle Obama}}_{\textit{Entity}_2}$ *on October 3, 1992.*

Spouse

Motivation

<USA, New York>, *LargestCity*

<USA, Washington D.C.>, *Capital*

<Richard Fuld, USA>, *Nationality*

Motivation

<USA, New York>, *LargestCity*
Capital: 0.5
LargestCity: 0.4
LocationCity: 0.05

<USA, Washington D.C.>, *Capital*
Capital: 0.95
LocationCity: 0.03

<Richard Fuld, USA>, *Nationality*
Nationality: 0.7
BirthPlace: 0.2

Motivation

<USA, New York>, *LargestCity*
Capital: 0.5
LargestCity: 0.4
LocationCity: 0.05

<USA, Washington D.C.>, *Capital*
Capital: 0.95
LocationCity: 0.03

<Richard Fuld, USA>, *Nationality*
Nationality: 0.7
BirthPlace: 0.2

Relation Constraints

Set	Sampled Positive Rules
C^{ts}	*(almaMater, knowFor), (city, region), (spouse, child)*
C^{to}	*(almaMater, owner), (city, hometown), (capital, city)*
C^{tso}	*(birthPlace, capital), (child, spouse), (city, country)*
C^{cs}	*almaMater, country, city, hometown*
C^{co}	*foundationPerson, child, knownFor, product*

Table 1: Example rules for each constraint set C^{ϕ}.

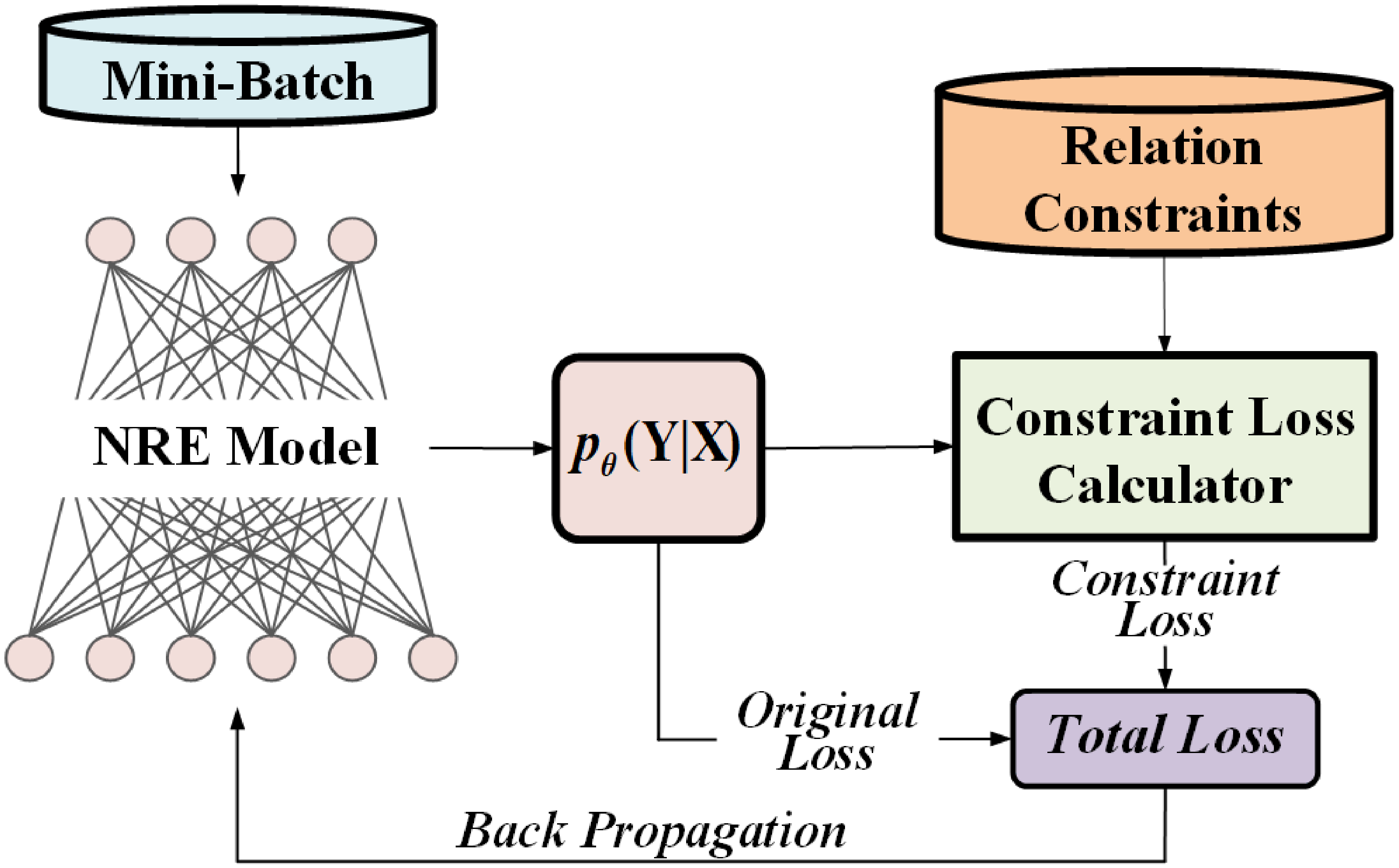
Relation Constraints

Set	Sampled Positive Rules
C^{ts}	*(almaMater, knowFor), (city, region),* $\boxed{\textit{(spouse, child)}}$
C^{to}	*(almaMater, owner), (city, hometown), (capital, city)*
C^{tso}	*(birthPlace, capital), (child, spouse), (city, country)*
C^{cs}	*almaMater, country, city, hometown*
C^{co}	*foundationPerson, child, knownFor, product*

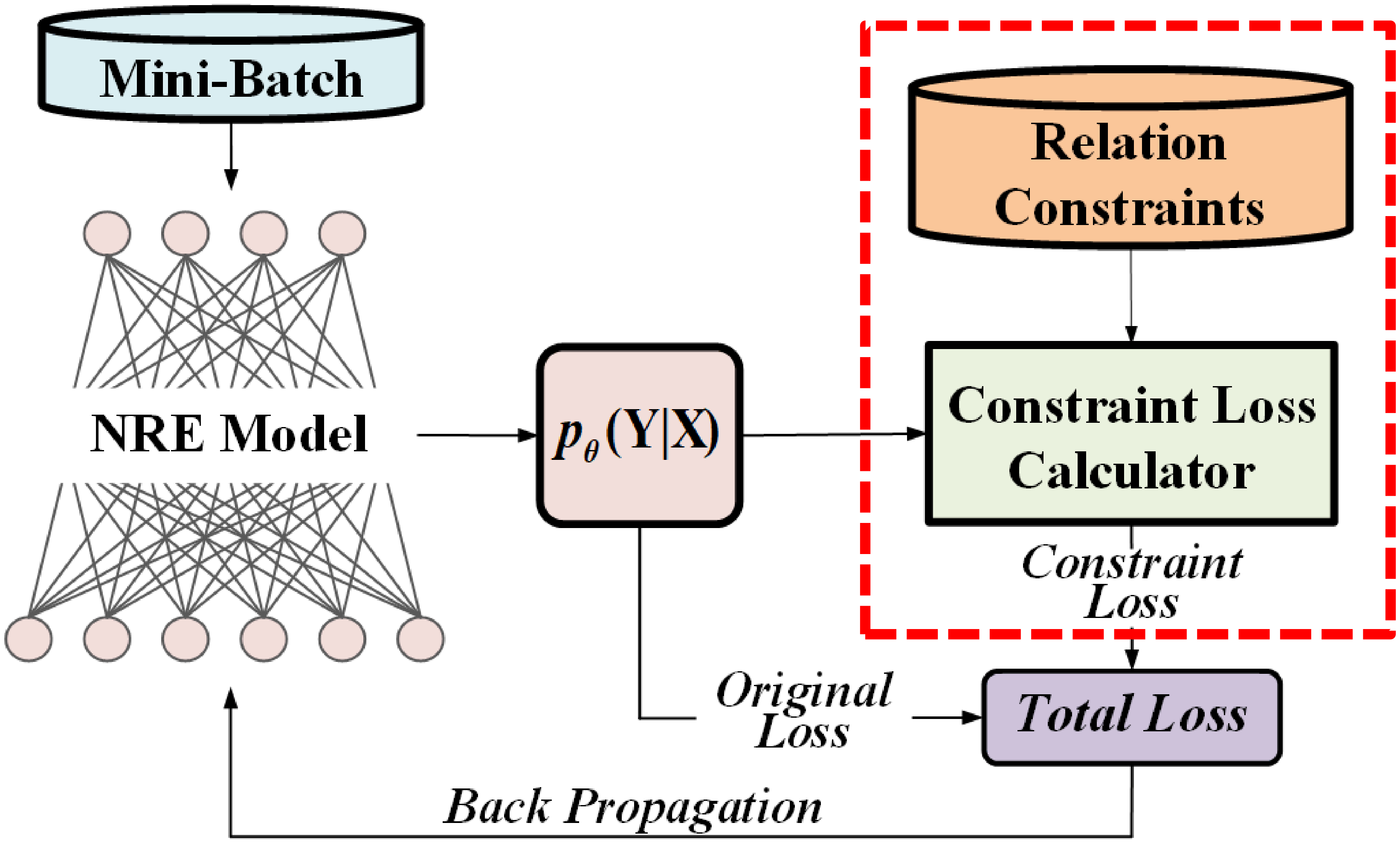
Table 1: Example rules for each constraint set C^{ϕ}.

Our Approach

北京大學 PEKING UNIVERSITY



Our Approach



Our Approach

北京大學
PEKING UNIVERSITY

A

Coherent

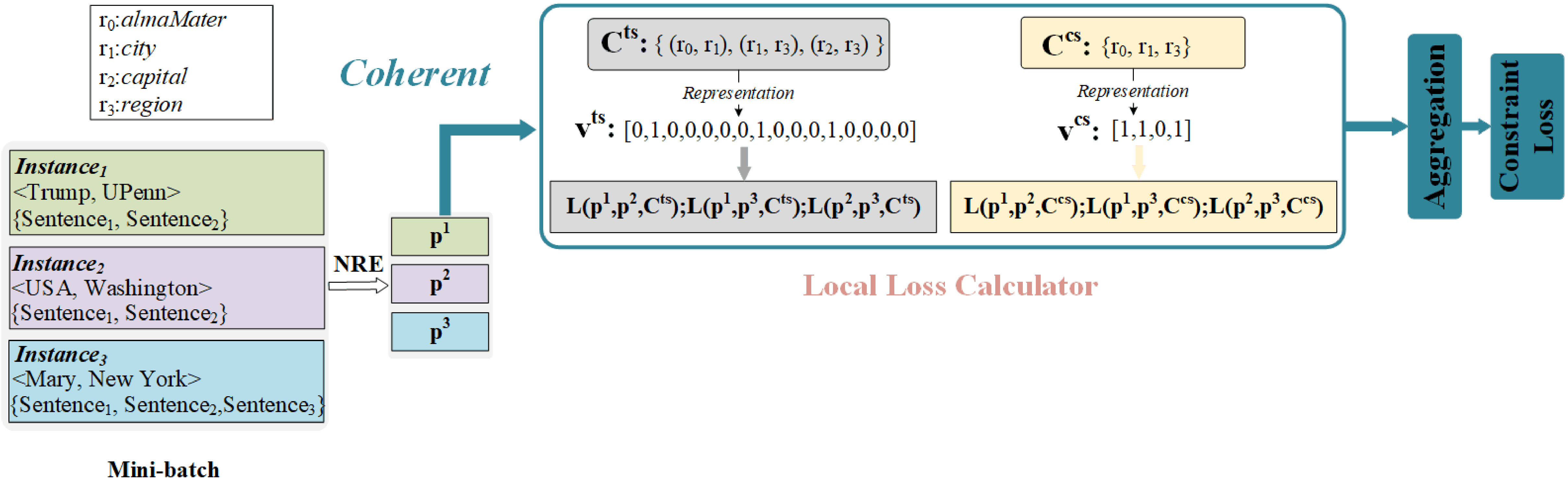
B

Semantic

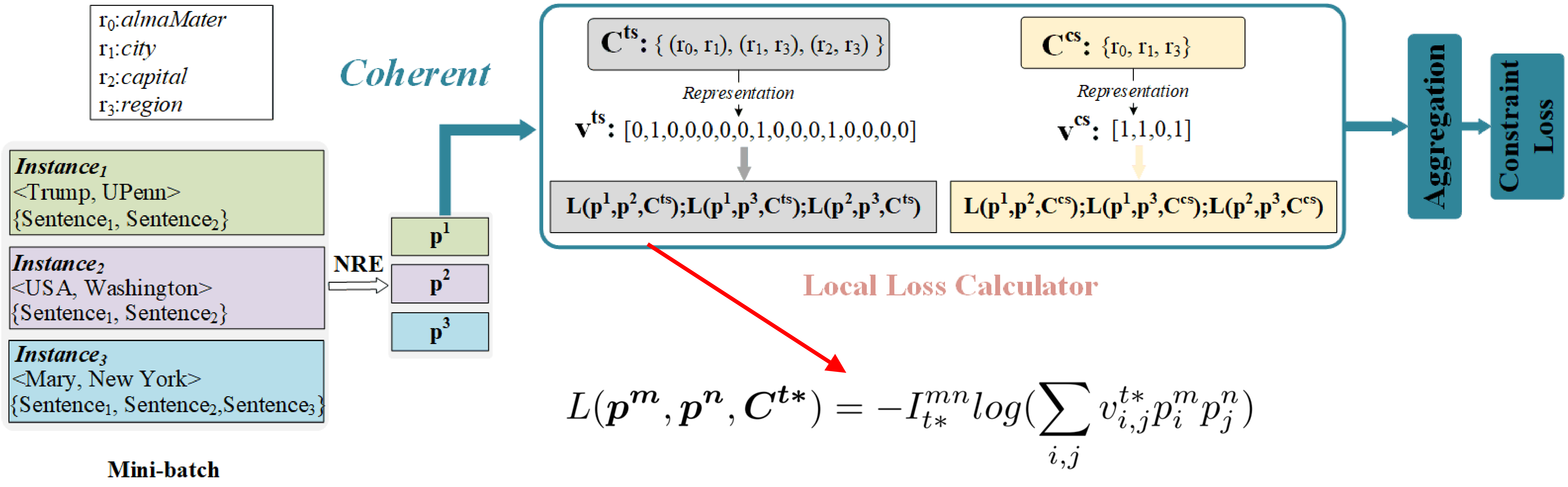
Coherent treats one constraint set as a whole.

Semantic pays more attention to which specific rules in the constraint sets the pairwise local predictions should satisfy.

Our Approach

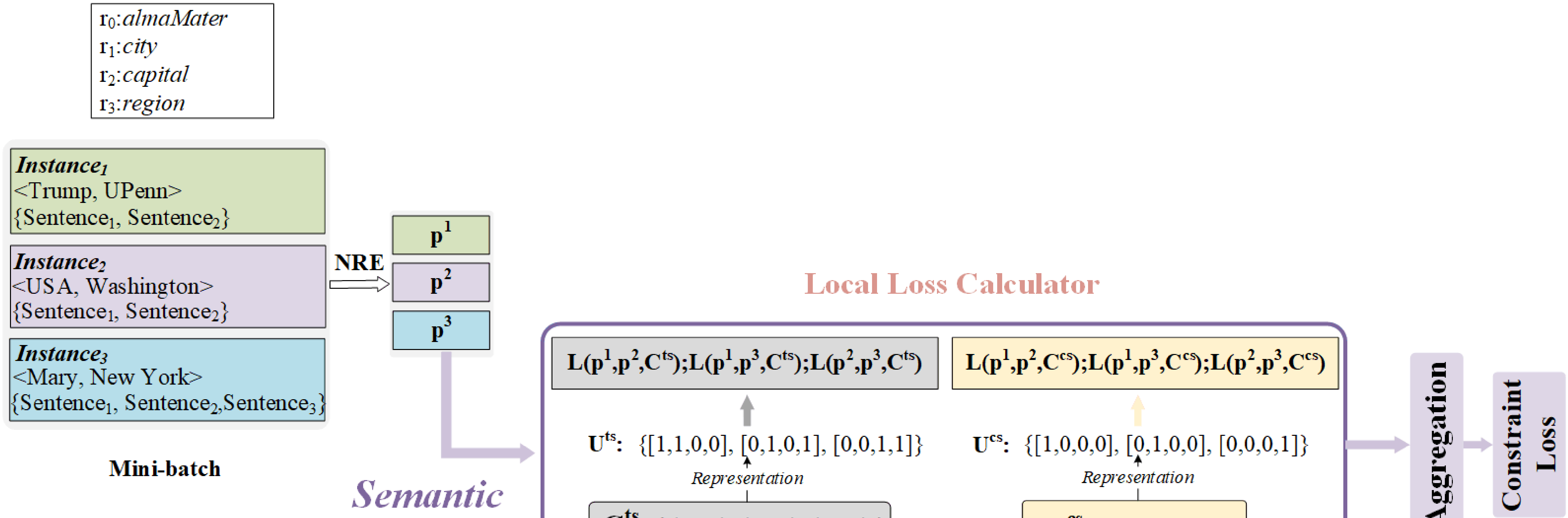


Our Approach



$$L(\boldsymbol{p^m}, \boldsymbol{p^n}, \boldsymbol{C^{t*}}) = -I_{t*}^{mn} log(\sum_{i,j} v_{i,j}^{t*} p_i^m p_j^n)$$

Our Approach

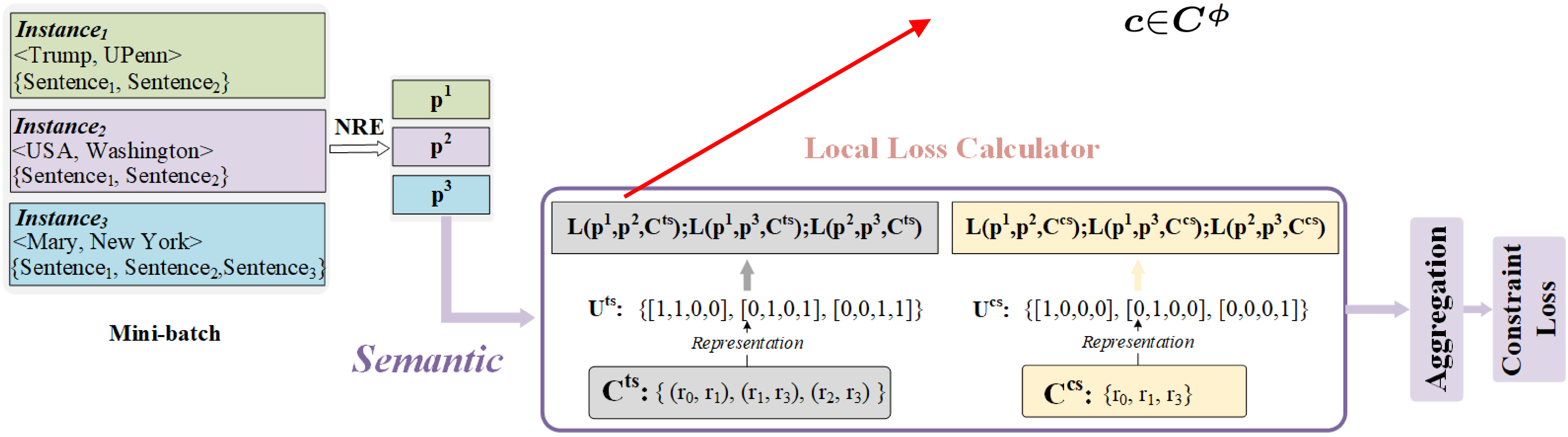


Our Approach

$$f(\boldsymbol{p^m}, \boldsymbol{p^n}, \boldsymbol{c}) = \prod_{u_i=1} q_i \prod_{u_i=0} (1 - q_i) \qquad q_i = p_i^m + p_i^n - p_i^m p_i^n$$

$$L(\boldsymbol{p^m}, \boldsymbol{p^n}, \boldsymbol{C^\phi}) = -I_\phi^{mn} log \sum_{\boldsymbol{c} \in \boldsymbol{C^\phi}} f(\boldsymbol{p^m}, \boldsymbol{p^n}, \boldsymbol{c})$$

r_0:*almaMater*
r_1:*city*
r_2:*capital*
r_3:*region*



Experiments

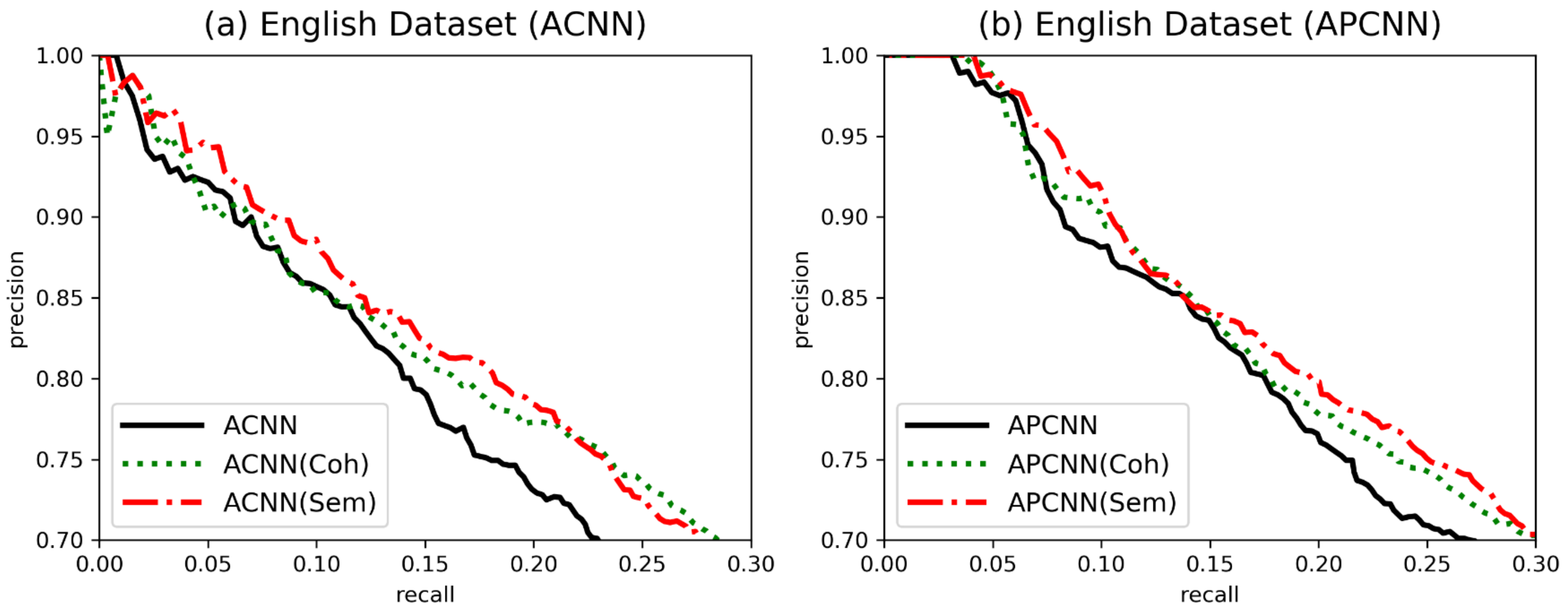
Datasets

- **English Dataset:** constructed by mapping triples in *Dbpedia* to sentences in the *New York Times* corpus
- **Chinese Dataset:** built by mapping the triples of *HudongBaiKe*, a large Chinese encyclopedia, with four Chinese economic newspapers

	Relations	Train Set		Test Set		Constraints
		# Triples	# Sent	# Triples	# Sent	
English Dataset	51	50k	134k	30k	53k	541
Chinese Dataset	28	60k	120k	40k	83k	110

Experiments

Main Results

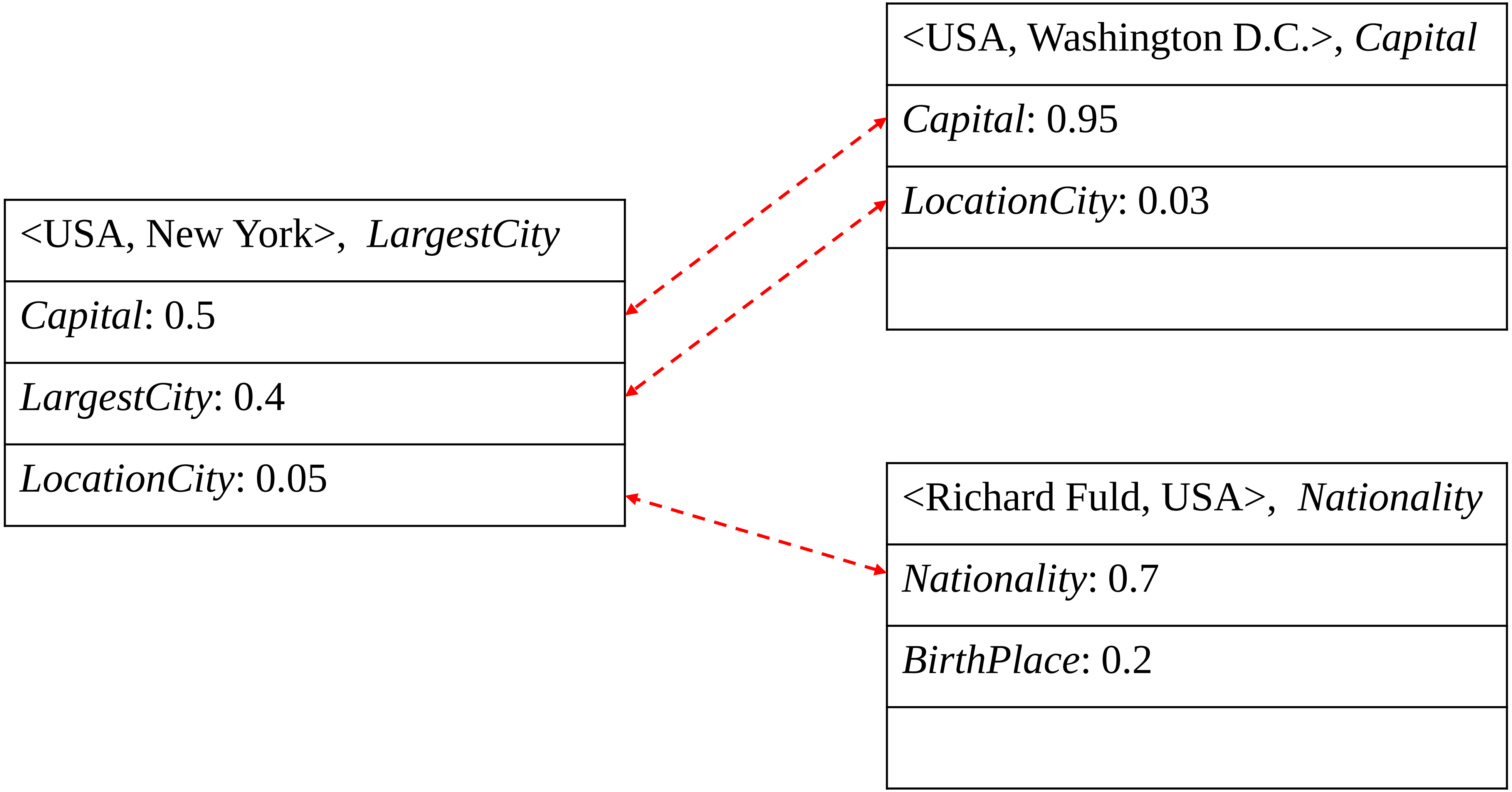


Experiments

北京大学
PEKING UNIVERSITY

Main Results

Model Name	English Dataset P@100	P@200	P@300	Mean	Δ_{Base}
ACNN	96.70	92.61	91.72	93.68	–
ACNN(Coh)	97.39	93.78	90.69	93.96	+0.3
ACNN(Sem)	97.62	95.87	94.12	95.87	+2.2
ACNN+ILP	97.87	94.36	93.16	95.13	+1.5
ACNN(Coh)+ILP	97.73	94.51	91.29	94.51	+0.8
ACNN(Sem)+ILP	**98.17**	**96.6**	**95.48**	**96.75**	**+3.1**
APCNN	100	98.97	97.41	98.79	–
APCNN(Coh)	100	99.57	97.33	98.97	+0.2
APCNN(Sem)	100	100	97.95	99.32	+0.5
APCNN+ILP	100	99.13	97.55	98.89	+0.1
APCNN(Coh)+ILP	100	100	98.03	99.34	+0.6
APCNN(Sem)+ILP	**100**	**100**	**98.39**	**99.46**	**+0.7**


<USA, New York>, *LargestCity*
Capital: 0.5
LargestCity: 0.4
LocationCity: 0.05
<USA, Washington D.C.>, *Capital*
Capital: 0.95
LocationCity: 0.03
<Richard Fuld, USA>, *Nationality*
Nationality: 0.7
BirthPlace: 0.2

Experiments

北京大學 PEKING UNIVERSITY

Main Results

Model Name	English Dataset P@100	P@200	P@300	Mean	Δ_{Base}
ACNN	96.70	92.61	91.72	93.68	–
ACNN(Coh)	97.39	93.78	90.69	93.96	+0.3
ACNN(Sem)	97.62	95.87	94.12	95.87	+2.2
ACNN+ILP	97.87	94.36	93.16	95.13	+1.5
ACNN(Coh)+ILP	97.73	94.51	91.29	94.51	+0.8
ACNN(Sem)+ILP	**98.17**	**96.6**	**95.48**	**96.75**	**+3.1**
APCNN	100	98.97	97.41	98.79	–
APCNN(Coh)	100	99.57	97.33	98.97	+0.2
APCNN(Sem)	100	100	97.95	99.32	+0.5
APCNN+ILP	100	99.13	97.55	98.89	+0.1
APCNN(Coh)+ILP	100	100	98.03	99.34	+0.6
APCNN(Sem)+ILP	**100**	**100**	**98.39**	**99.46**	**+0.7**

After equipped with our *CLC* modules, i.e., *Coherent* and *Semantic*, both *ACNN* and *APCNN* obtain improvement on the English dataset.

Experiments

Main Results

Model Name	English Dataset P@100	P@200	P@300	Mean	Δ_{Base}
ACNN	96.70	92.61	91.72	93.68	–
ACNN(Coh)	97.39	93.78	90.69	93.96	+0.3
ACNN(Sem)	97.62	95.87	94.12	95.87	+2.2
ACNN+ILP	97.87	94.36	93.16	95.13	+1.5
ACNN(Coh)+ILP	97.73	94.51	91.29	94.51	+0.8
ACNN(Sem)+ILP	**98.17**	**96.6**	**95.48**	**96.75**	**+3.1**
APCNN	100	98.97	97.41	98.79	–
APCNN(Coh)	100	99.57	97.33	98.97	+0.2
APCNN(Sem)	100	100	97.95	99.32	+0.5
APCNN+ILP	100	99.13	97.55	98.89	+0.1
APCNN(Coh)+ILP	100	100	98.03	99.34	+0.6
APCNN(Sem)+ILP	**100**	**100**	**98.39**	**99.46**	**+0.7**

Applying ILP to our approach can obtain further improvement.

Experiments

北京大學 PEKING UNIVERSITY

Main Results

Model Name	English Dataset P@100	P@200	P@300	Mean	Δ_{Base}
ACNN	96.70	92.61	91.72	93.68	–
ACNN(Coh)	97.39	93.78	90.69	93.96	+0.3
ACNN(Sem)	97.62	95.87	94.12	95.87	+2.2
ACNN+ILP	97.87	94.36	93.16	95.13	+1.5
ACNN(Coh)+ILP	97.73	94.51	91.29	94.51	+0.8
ACNN(Sem)+ILP	**98.17**	**96.6**	**95.48**	**96.75**	**+3.1**

1.5

0.5

0.9

Conclusions

- A unified framework to integrate relation constraints with neural networks for RE
  - Coherent (general perspectives), Semantic (precise perspectives)
  - Validating our approach on English and Chinese datasets
- Our study reveals that learning with the constraints can better utilize the constraints from a different perspective compared to the ILP post-processing method

Thank you!

Q&A

pkuyeyuan@pku.edu.cn

WICT, PKU