

Probabilistic Reasoning for Entity & Relation Recognition

Seminar für Computerlinguistik

Logik in der Praxis

Lyubov Nakryyko



Probabilistic Reasoning for Entity and Relation Recognition

- Einführung
- Globale Entitäten/
Relationeninferenz
- Berechnungsverfahren
- Experimente
- Verbesserungsvorschläge

Einführung

Erkennung und Klassifikation von Entitäten und Relationen in Korpora ist die Hauptaufgabe in vielen NLP Anwendungen:

- Informationsextraktion(IE)
 - Question Answering(QA)
 - Textzusammenfassung
- „Where was Poe born?“(Ortsentität?)
- „Who killed Lee Harvey Oswald?“(Entität einer Person, die Lee Harvey Oswald umgebracht hat?)

Alter Ansatz

zuerst: Identifizierung und Klassifizierung der Entitäten

dann: Ermitteln der Relationen zwischen den Entitäten

Probleme:

- falsche Klassifizierung der Entitäten
- ambige Entitätsnamen

Neuer Ansatz

ein probabilistisches System , das die Entitäten und Relationen gleichzeitig erkennt

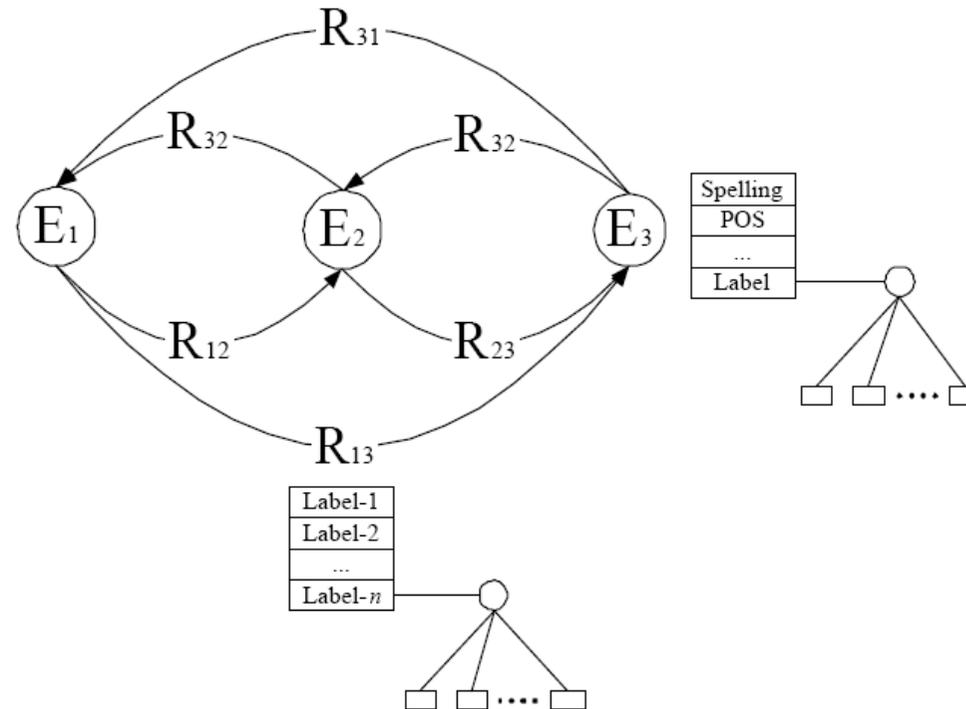
- ✓ Dabei werden die Entitäts- und Relationsklassen erstellt in Bezug auf bedingte Distrubution jeder Entität/ Relation gegeben beobachtete Daten
- ✓ constraints werden Relationen und Entitäten zugeordnet

Beispiel: (das erste Argument von *kill* ist sehr wahrscheinlich *person*, das zweite Argument von *born_in* ist *location*)

Globale Entitäten/

Relationeninferenz

Das Problem zu lösen: satzweise kohärente Klassifizierung aller Entitäten und (binärer) Relatio



Entitäten und Relationen mit den beiderseitigen Abhängigkeiten können als markierter Graph dargestellt werden

Globale Entitäten/Relationeninferenz

Beispiel:

Dieser Satz

“Dole ’s wife , Elizabeth , is a native of Salisbury , N.C.” hat drei Entitäten:

E1 = “Dole”, E2 = “Elizabeth”, and E3 = “Salisbury, N.C.”

Dole	's wife ,	Elizabeth	, is a native of	Salisbury , N.C.
E1		E2		E3

Global Inference of Entities/

Relations

Der gleiche Satz beinhaltet folgende Relationen zwischen den Entitäten:

$R_{12} = (\text{"Dole"}, \text{"Elizabeth"})$

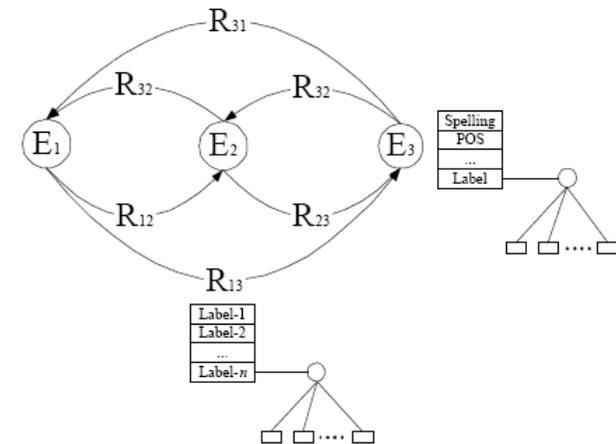
$R_{21} = (\text{"Elizabeth"}, \text{"Dole"})$

$R_{13} = (\text{"Dole"}, \text{"Salisbury, N.C."})$

$R_{31} = (\text{"Salisbury, N.C."}, \text{"Dole"})$

$R_{23} = (\text{"Elizabeth"}, \text{"Salisbury, N.C."})$

und $R_{32} = (\text{"Salisbury, N.C."}, \text{"Elizabeth"})$



Global Inference of Entities/

Relations

Menge der Entitätsklassen CE = {other_ent, person, location}

Menge der Relationsklassen CR = {other_rel, born_in, spouse_of}

E1, E2 gehört zu person

E3 gehört zu location

R23 ist born_in

R12, R21 sind spouse_of

Andere Relationen sind other_rel

Global Inference of Entities/Relations

Ein Constraint C ist ein 3-Tupel $(R; \varepsilon_1; \varepsilon_2)$, mit $R \in CR$ and $\varepsilon_1; \varepsilon_2 \in CE$. Wenn eine Relation das Label R hat, dann sind die erlaubten Klassenlabels für ihrer Argumente zwei Entitäten ε_1 und ε_2 .

Beispiel: $(born_in, person, location)$,
 $(spouse_of, person, person)$ und $(murder, person, person)$

Eigenschaft 1 Die

Labels Wahrscheinlichkeit *einer relation* R_{ij} gegeben die Labels ihrer Argumente E_i und E_j hat folgende Eigenschaften:

- $P(R_{ij} = \text{other_rel} | E_i = e_1, E_j = e_2) = 1$, wenn kein r existiert, sodass $(r; e_1; e_2)$ ein Constraint ist
- $P(R_{ij} = r | E_i = e_1, E_j = e_2) = 0$, wenn es kein Constraint c gibt, so dass $c = (r; e_1; e_2)$.



Berechnungsverfahren

- Die Klassenlabels von Entitäten und Relationen hängen von vielen Variablen ab
- Um die wahrscheinlichsten kohärenten Labels vorauszusagen werden Inferenzregeln auf die Variablen angewendet
- Bei der Modellierung der Interaktion zwischen Zieleigenschaften werden die bedingten Wahrscheinlichkeiten der Beobachtungen benutzt

Basic Klassifikator

Genauere Voraussage dieser zwei Klassifikatoren ist auf syntaktische und semantische Analyse des ganzen Satzes basiert

„**Schwache**“ **Klassifikatoren** werden aus den lokalen Eigenschaften der Wörter in oder um den Zielbereich ermittelt:

- Schreibweise des Wortes
- Wortklasse
- Semantische Information aus externe Quellen wie z.B. WordNet

Für diesen Zweck kann SNoW(multi-class classifier) benutzt werden

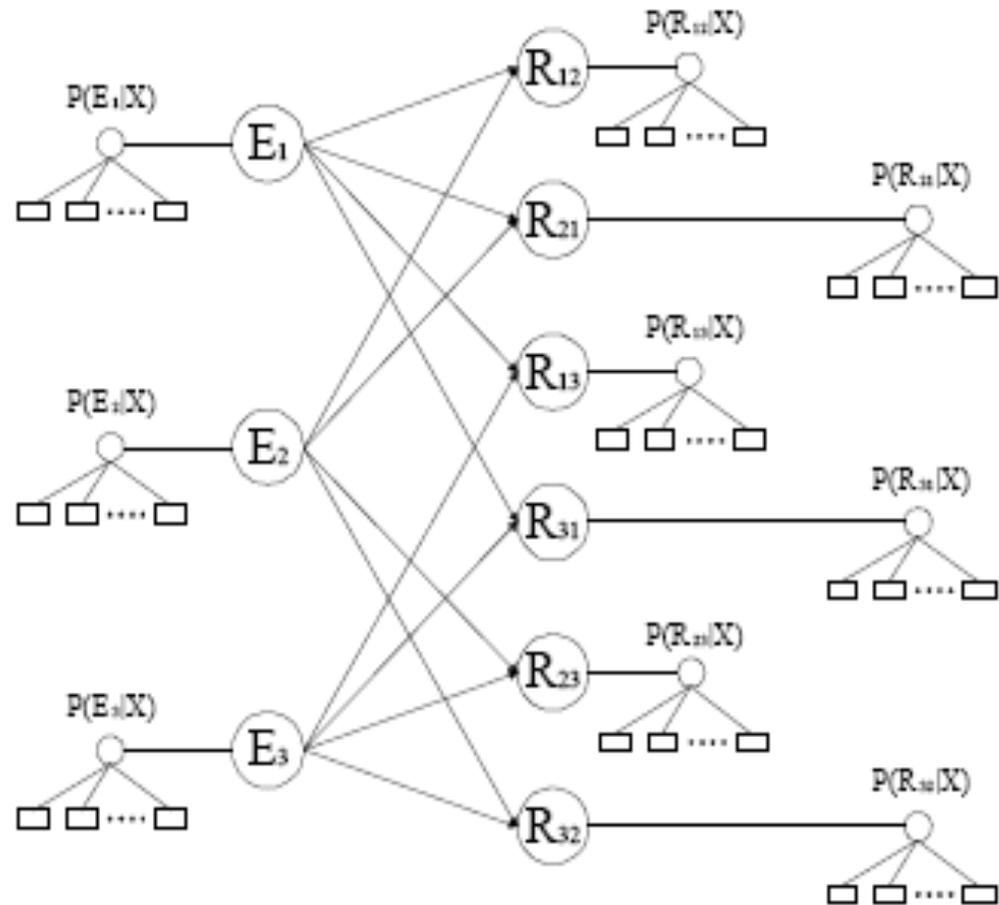
Bayessches Inferenz Modell

Belief Network

$P(E|Beobachtungen)$

$P(R|Beobachtungen)$

$P(R_{ij}|E_i, E_j)$



Experimente

Datenmaterial(Wall Street Journal, Associated Press, und San Jose Mercury News)

- 245 Sätze mit *kill*(*murder-victim* Relation)
- 176 Sätze mit *born_in* Relation
- und 502 Sätze, die keine Relationen beinhalten

Vordefinierte Klassenlabels:

other_ent, person, location, other_rel, kill and birthplace

Drei Datensätze werden erstellt: „kill“, „born_in“ und „alles“



Testansätze

- Basic (Ausführung des basic Klassifikators)
- Omniscient
- BN(Bayes Net Toolbox for Matlab by Murphy)

Resultaten

BN (recall -, precision +, F1 of relation classification +, F1 of entities classification -)

Approach	person			location		
	Rec	Prec	F_1	Rec	Prec	F_1
Basic	96.6	92.3	94.4	76.3	91.9	83.1
BN	89.0	96.1	92.4	78.8	86.3	82.1
Omniscient	96.4	92.6	94.5	75.4	90.2	81.9

Approach	kill		
	Rec	Prec	F_1
Basic	61.8	57.2	58.6
BN	49.8	85.4	62.2
Omniscient	67.7	63.6	64.8

Approach	person			location		
	Rec	Prec	F_1	Rec	Prec	F_1
Basic	85.5	90.7	87.8	89.5	93.2	91.1
BN	87.0	90.9	88.8	87.5	93.4	90.3
Omniscient	90.6	93.4	91.7	90.7	96.5	93.4

Approach	born_in		
	Rec	Prec	F_1
Basic	81.4	63.4	70.9
BN	87.6	70.7	78.0
Omniscient	86.9	71.8	78.0

Approach	person			location		
	Rec	Prec	F_1	Rec	Prec	F_1
Basic	92.1	87.0	89.4	83.2	81.1	82.0
BN	78.8	94.7	86.0	83.0	81.3	82.1
Omniscient	93.4	87.3	90.2	83.5	83.1	83.2

Approach	kill			born_in		
	Rec	Prec	F_1	Rec	Prec	F_1
Basic	43.8	78.6	55.0	69.0	72.9	70.5
BN	47.2	86.8	60.7	68.4	87.5	76.6
Omniscient	52.8	79.5	62.1	76.1	71.3	73.2



Zukünftige Pläne

- Boot-strapping
- Verbesserung Annäherungen bei der Berechnung
- Belief network Modell hat nicht genug Ausdrucksfähigkeit