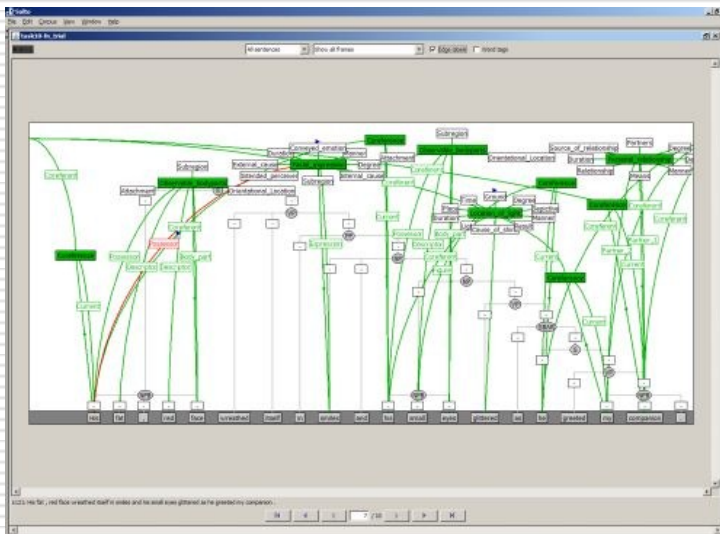


Linking Events and Their Participants in Discourse



Lyubov Nakryyko
Eric Hildebrand
Mateusz Dolata

Softwareprojekt WS09/10

Ruprecht-Karls-Universität Heidelberg

Seminar für Computerlinguistik

Institut für allgemeine und angewandte
Sprach- und Kulturwissenschaft

Inhaltsverzeichnis

- SemEval-2010 Task #10
 - Überblick
 - Exkurs: SLR
 - Exkurs: NIs
 - Unser Subtask
- Geplantes Vorgehen
 - Architektur (Skizze)
 - a) Core Roles
 - b) Licensors & Interpretation
 - c) Role Linking

Inhaltsverzeichnis

☐ SemEval-2010 Task #10

- Überblick
- Exkurs: SLR
- Exkurs: NIs
- Unser Subtask

☐ Geplantes Vorgehen

- Architektur (Skizze)
- a) Core Roles
- b) Licensors & Interpretation
- c) Role Linking

Inhaltsverzeichnis

- SemEval-2010 Task #10
 - Überblick
 - Exkurs: SLR
 - Exkurs: NIs
 - Unser Subtask
- Geplantes Vorgehen
 - Architektur (Skizze)
 - a) Core Roles
 - b) Licensors & Interpretation
 - c) Role Linking

SemEval-2010 Task #10

Überblick

□ Variante von SRL

- neu: linking lokaler semantischer Argumentstrukturen (LSA) über Satzgrenzen hinaus
- genauer: lokale null-instantiierte Rollen mit ihren Koreferenten in umgebenden Sätzen verbinden

Inhaltsverzeichnis

- SemEval-2010 Task #10
 - Überblick
 - Exkurs: SLR
 - Exkurs: NIs
 - Unser Subtask
- Geplantes Vorgehen
 - Architektur (Skizze)
 - a) Core Roles
 - b) Licensors & Interpretation
 - c) Role Linking

SemEval-2010 Task #10

Exkurs: SRL (1)

- Semantic Role Labelling
 - Zuweisung semantischer Rollen zu syntaktischen Argumenten eines Prädikats
 - SR'n beschreiben Funktionen von Diskurspartizipanten (*wer hat was wem wie, wann, wo,... getan?*)

SemEval-2010 Task #10

Exkurs: SRL (2)

- Semantic Role Labelling
 - Problem: bisherige Ansätze traditionell lokal (satzintern)
 - Beziehungen zwischen verschiedenen lokalen LSAs werden ignoriert
 - ⇒ Verlust wichtiger semantischer Informationen

SemEval-2010 Task #10

wir erinnern uns...

„It is clear that there is an interplay between local argument structure and the surrounding discourse.“

(Ruppenhofer et al. 2009, acc. to Fillmore 1977)

SemEval-2010 Task #10

Beispiele (1)

- (1) In a lengthy court case the defendant was tried for murder. In the end, he was cleared.
- (2) Watson won't allow that I know anything of art but that is mere jealousy because our views upon the subject differ.
- (3) Twenty-two month old with history of recurrent right middle lobe infiltrate. Increased cough, tachypnea, and work of breathing.

Inhaltsverzeichnis

☐ SemEval-2010 Task #10

- Überblick

- Exkurs: SLR

- Exkurs: NIs

- Unser Subtask

☐ Geplantes Vorgehen

- Architektur (Skizze)

- a) Core Roles

- b) Licensors & Interpretation

- c) Role Linking

SemEval-2010 Task #10

Exkurs: NIs (1)

- Null-Instantiierungen
 - Auslassung von Core-Argumenten eines Prädikats
- Kategorisierung des Phänomens durch
 - Licensor
 - Interpretation

SemEval-2010 Task #10

Exkurs: NIs (2)

- Licensor: Auslassung nur, wenn
 - a) ein bestimmtes lexikalisches Element vorhanden ist oder
 - b) eine bestimmte grammatikalische Konstruktion vorliegt
- Interpretation
 - definit oder
 - indefinit

SemEval-2010 Task #10

Beispiele (2)

(1) No doubt, mistakes were made
O^{Protagonist}.

Licenser: Passiv (grammatikalisch)

Interpretation: indefinit

(2) We arrived *O^{Goal}* at 8pm.

Licenser: *arrived* (lexikalisch)

Interpretation: definit

Inhaltsverzeichnis

☐ SemEval-2010 Task #10

- Überblick
- Exkurs: SLR
- Exkurs: NIs
- Unser Subtask

☐ Geplantes Vorgehen

- Architektur (Skizze)
- a) Core Roles
- b) Licensors & Interpretation
- c) Role Linking

SemEval-2010 Task #10

Unser Subtask (1)

□ Full Task

- (1) find the overt semantic arguments of the target (role recognition)
- (2) label them with correct role (role labelling)
- (3) recognize definite null instantiations and find links to antecedents in the wider context (NI linking)

□ NIs only

- entspricht Full Task (3)
- ⇒ unsere Aufgabe

SemEval-2010 Task #10

Unser Subtask (2)

NI linking

- a) Recognize that a core role is missing
- b) Ascertain that it should have a definite filler (Identify Definite Null Instantiations and associated Licensor)
- c) Find an antecedent for it (in the discourse)

Inhaltsverzeichnis

☐ SemEval-2010 Task #10

- Überblick
- Exkurs: SLR
- Exkurs: NIs
- Unser Subtask

☐ Geplantes Vorgehen

- Architektur (Skizze)
 - a) Core Roles
 - b) Filler
 - i. Interpretation
 - ii. Licensor
 - c) Role Linking

Inhaltsverzeichnis

☐ SemEval-2010 Task #10

- Überblick
- Exkurs: SLR
- Exkurs: NIs
- Unser Subtask

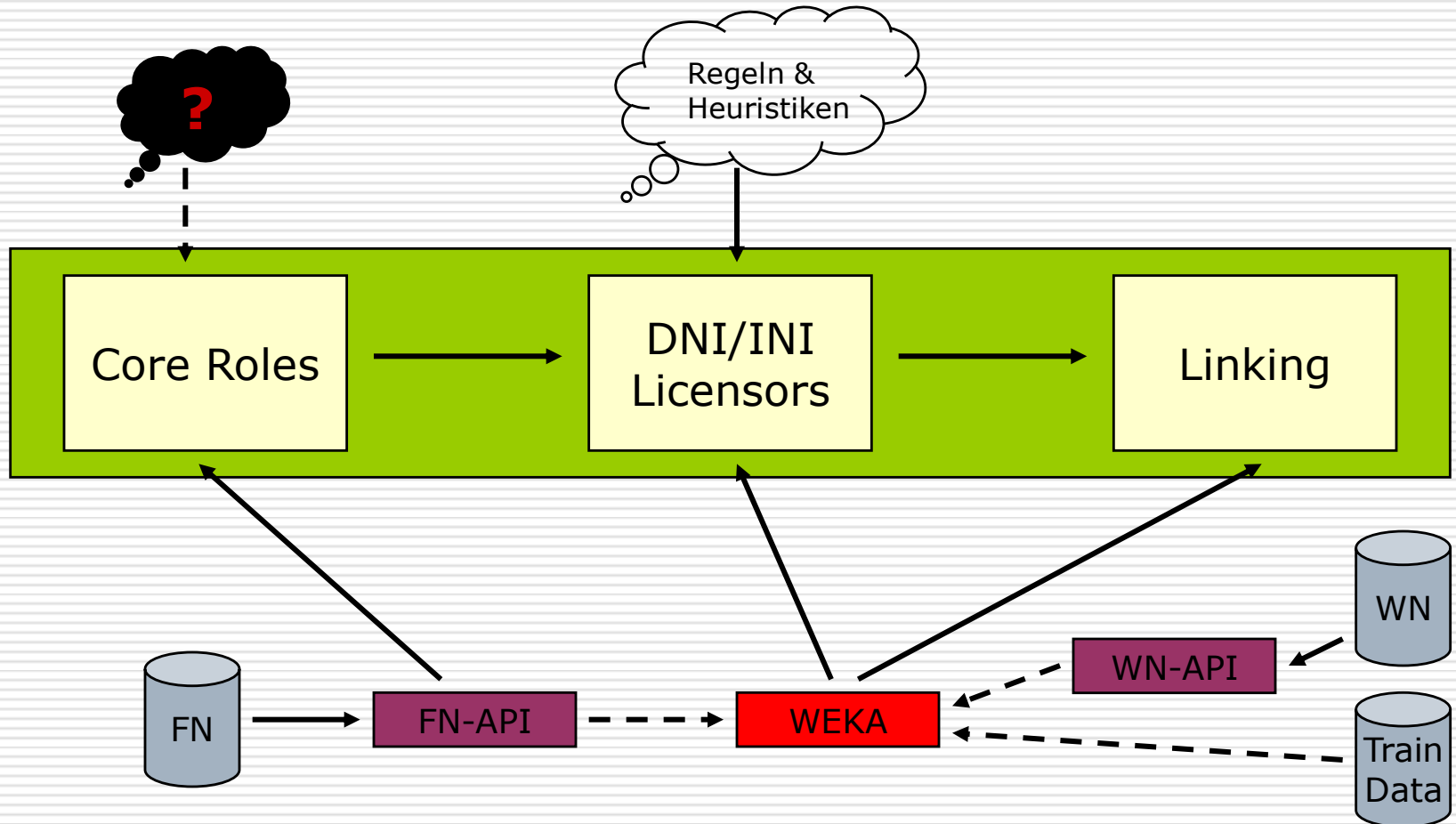
☒ Geplantes Vorgehen

■ Architektur (Skizze)

- a) Core Roles
- b) Filler
 - i. Interpretation
 - ii. Licensor
- c) Role Linking

SemEval-2010 Task #10

Architektur (Skizze)



Inhaltsverzeichnis

☐ SemEval-2010 Task #10

- Überblick
- Exkurs: SLR
- Exkurs: NIs
- Unser Subtask

☒ Geplantes Vorgehen

- Architektur (Skizze)
 - a) Core Roles
 - b) Filler
 - i. Interpretation
 - ii. Licensor
 - c) Role Linking

SemEval-2010 Task #10

Unser Subtask (2)

NIs only

- a) Recognize that a core role is missing
- b) Ascertain that it should have a definite filler (Identify Definite Null Instantiations and associated Licensor)
- c) Find an antecedent for it (in the discourse)

SemEval-2010 Task #10

a) Core Roles (1)

„First, one needs to determine whether a core role is missing. This involves looking up which core roles are overtly expressed and which are not.“

(Ruppenhofer et al. 2009)

Klingt einfach, aber...

SemEval-2010 Task #10

a) Core Roles (2)

- Annotation der Trail-Frames nicht übereinstimmend mit online FN-Daten (sehr oft einfach alle Rollen Core Roles)
- nicht jede nicht-overt Core-Role muss zwangsweise im Diskurs realisiert sein - manche *FEs stehen in Relation zueinander...*

SemEval-2010 Task #10

a) FE-Relationen (1)

1. Coreness-Sets

- in vielen Frames *genügt es der semantischen Valenz, wenn eine der Rollen aus einem Core-Set realisiert ist*
- schließt nicht das Vorkommen mehrerer aber aus

SemEval-2010 Task #10

a) FE-Relationen (3)

2. Requires

- in manchen Fällen *erzwingt* das Vorkommen eines bestimmten Frameelements die *das Auftreten eines anderen*, sonst ungrammatisch
- auch Non-Core Roles können plötzlich erforderlich werden
- syntaktisch, für uns weniger interessant

SemEval-2010 Task #10

a) FE-Relationen (3)

3. Excludes

- ebenso *schließt* in anderen Fällen das Vorkommen eines bestimmten Frameelements *das Auftreten von anderen aus*
- in diesen Fällen dürfen wir *die ausgeschlossenen Rollen nicht suchen*

SemEval-2010 Task #10

a) Core Roles (3)

- Dieser Informationen sind in FrameNet nicht vollständig kodiert, die Annotationsbeispiele nicht exhaustiv
- in der Task-Description werden sie nicht erwähnt - ignorieren?
 - dann Lösung trivial, kann direkt aus der Frame-Definition abgelesen und mit Satz verglichen werden

Inhaltsverzeichnis

☐ SemEval-2010 Task #10

- Überblick
- Exkurs: SLR
- Exkurs: NIs
- Unser Subtask

☐ Geplantes Vorgehen

■ Architektur (Skizze)

a) Core Roles

b) Filler

i. Interpretation

ii. Licensor

c) Role Linking

SemEval-2010 Task #10

Unser Subtask (2)

NIs only

- a) Recognize that a core role is missing
- b) Ascertain that it should have a definite filler (Identify Definite Null Instantiations and associated Licensor)
- c) Find an antecedent for it (in the discourse)

SemEval-2010 Task #10

b) Licensors & Interpretation (1)

- Licensors bestimmt eindeutig Interpretation
- beides in den Trainingsdaten annotiert (bei Licensors nur lexical/constructional)
- „To do this, one can use rules and heuristics based on various syntactic and lexical facts of English“)

(Ruppenhofer et al. 2009)

SemEval-2010 Task #10

b) Licensors & Interpretation (2)

„As for using lexical knowledge, consider omissions of the Goods FE of the verb steal in the Theft frame. FrameNet annotation shows that whenever the Goods FE of steal is missing it is interpreted indefinitely, suggesting that a new instance of the FE being missing should have the same interpretation. “

(Ruppenhofer et al. 2009)

Wirklich?...

SemEval-2010 Task #10

b) Licensors & Interpretation (3)

- ❑ Aussage trifft nicht allgemein auf die FN-Annotationen zu
- ❑ oftmals mehrere Interpretationen für Auslassung derselben Rolle, Licensors in FN nicht annotiert
- ⇒ Lernen daher nur aus eindeutigen und umfangreichen Beispielen möglich (wahrscheinlich nur sehr wenige)

Inhaltsverzeichnis

☐ SemEval-2010 Task #10

- Überblick
- Exkurs: SLR
- Exkurs: NIs
- Unser Subtask

☐ Geplantes Vorgehen

- Architektur (Skizze)
 - a) Core Roles
 - b) Filler
 - i. Interpretation
 - ii. Licensor
 - c) Role Linking

SemEval-2010 Task #10

i) Interpretation - Features (1)

- unmittelbar zugänglich (aus Annotation der Trainingsdaten):
 - FrameName
 - FE-Name
 - POS-Tag
 - übergeordnete Phrasentyp(en)

SemEval-2010 Task #10

i) Interpretation - Features (2)

□ zusätzlich

- [Synset:ID] (WN-Synset des Targets (eindeutige Bestimmung durch Prüfung, ob weitere potentielle Targets (in FN aufgelistet) Teil des Synsets sind)
- [possesive:+/-] (beginnt NÄCHST-übergeordnete Nominalphrase mit PossPron (feste Liste) oder <Nomen>'s)

SemEval-2010 Task #10

i) Interpretation - Features (3)

- zusätzlich

- [SimSet:ID] (FN-SimSet: fassen „ähnliche“ Frames in Set zusammen)
 - die Frames, bei denen ein bestimmter Anteil von Rollennamen übereinstimmt

SemEval-2010 Task #10

i) Interpretation - Features (4)

☐ zusätzlich

■ wenn N:

- ☐ [def:+/-] (Kopf definiter NP?) [aufsteigen bis 1. NP Knoten, erstes Wort der Phrase prüfen]
- ☐ [pl:+/-] (Plural/Singular)

■ wenn V:

- ☐ [inf:+/-] (Infinitiv?)
- ☐ [subj:+/-] (hat Subjekt?)
- ☐ wenn [subj: +]: [synSubj:ID] (Synset-ID) des Subjekts

Inhaltsverzeichnis

☐ SemEval-2010 Task #10

- Überblick
- Exkurs: SLR
- Exkurs: NIs
- Unser Subtask

☐ Geplantes Vorgehen

- Architektur (Skizze)
 - a) Core Roles
 - b) Filler
 - i. Interpretation
 - ii. Licensor
 - c) Role Linking

SemEval-2010 Task #10

ii) Licensor - Features (1)

- unmittelbar zugänglich (aus Annotation der Trainingsdaten):
 - FrameName
 - FE-Name
 - POS-Tag
 - übergeordnete Phrasentyp(en)

SemEval-2010 Task #10

ii) Licensor - Features (2)

□ zusätzlich

- [Synset:ID] (WN-Synset des Targets (eindeutige Bestimmung durch Prüfung, ob weitere potentielle Targets (in FN aufgelistet) Teil des Synsets sind)
- [possessive:+/-] (beginnt NÄCHST-übergeordnete Nominalphrase mit PossPron (feste Liste) oder <Nomen>'s)
- [TargetWithinMessage:+/-] (Frame-evozierendes Token innerhalb eines „Message“/„Content“-FEs?)

SemEval-2010 Task #10

ii) Licensor - Features (3)

☐ zusätzlich

■ wenn N:

- ☐ [def:+/-] (Kopf definiter NP?)
[aufsteigen bis 1. NP Knoten, erstes Wort der Phrase prüfen]
- ☐ [pl:+/-] (Plural/Singular)
- ☐ [TargetWithinThereArePhrase:+/-]
(Frame-evozierendes Token innerhalb einer <There are ...>-Phrase?)

SemEval-2010 Task #10

ii) Licenser - Features (4)

- zusätzlich
 - wenn V:
 - [inf:+/-] (Infinitiv?)
 - [subj:+/-] (hat Subjekt?)
 - wenn [subj: +]: [synSubj:ID] (Synset-ID) des Subjekts
 - [TargetWithinCite:+/-] (Frame-evozierendes Token innerhalb wörtlicher Rede?)
 - [TargetHeadOfXComp:+/-] (Frame-evozierendes Token Kopf einer eingebetteten Infinitivstruktur)

Inhaltsverzeichnis

☐ SemEval-2010 Task #10

- Überblick
- Exkurs: SLR
- Exkurs: NIs
- Unser Subtask

☐ Geplantes Vorgehen

- Architektur (Skizze)
 - a) Core Roles
 - b) Filler
 - i. Interpretation
 - ii. Licensor
 - c) Role Linking

SemEval-2010 Task #10

Unser Subtask (2)

NIs only

- a) Recognize that a core role is missing
- b) Ascertain that it should have a definite filler (Identify Definite Null Instantiations and associated Licensor)
- c) Find an antecedent for it (in the discourse)

SemEval-2010 Task #10

c) Role Linking

- betrachten als Koreferenz-auflösungsproblem
- auch hier: Lernen mit Features für potentielle Role fillers, dann Vergleich mit allen im Diskurs in Frage kommenden
- Lernen aus Trainingsdaten + annotierten FN-Daten

SemEval-2010 Task #10

c) Antecedenten - Features (1)

- unmittelbar zugänglich (aus Annotation der Trainingsdaten):
 - FrameName
 - FE-Name
 - POS-Tag
 - übergeordnete Phrasentyp(en)

SemEval-2010 Task #10

c) Antezedenten - Features (2)

- zusätzlich
 - [strMatch] (exaktes Matching, aus overten Beispielen von FN LU-Annotation)
 - [AntSynset:ID] (aus overten Beispielen Synsets für Antezedenten finden)
 - [dist:n] (Abstand (in grammatikalischen Knoten, aus annotierten Trainingsbeispielen)
 - [sentDist:n] (Abstand bei satzübergreifenden Fillern)

SemEval-2010 Task #10

c) Antecedenten - Features (3)

☐ zusätzlich

- [AnaphoricDensity:n] (#satzübergreifender Koreferenzen (annotiert in Trainingsdaten))
- [FEOverlap:<FE-ID,F-ID>] (aus overtten Beispielen: Filler weiterem Frame F zugewiesen (weitere FE)? $\Rightarrow$ finde Targets im Diskurs, die die gleiche FE einnehmen (in „ähnlichen“ Frame (aus SimSet, s.o.))
- wenn N:
 - ☐ [pl:+/-] (Plural/Singular)
 - ☐ [NE-type:ID] (des Kopfs der übergeordneten NP; nur falls NE)

Abschluss

Vielen Dank für die
Aufmerksamkeit! 😊
Gibt es Fragen?