Einführung in die Pragmatik und Diskurs:

Diskurs 2: Computermodele für Referenz und Diskurs

M. Pinkal/M. Wolska/C. Sporleder

Universität des Saarlandes

Sommersemester 2007

Übersicht

Vorlesungsplan

- Referenzresolution
- Diskursinterpretation

Kernlektüre

- Jurafsky & Martin (2000), Kapitel 18

Teil 1: Referenzresolution

Referenzresolution

Der Briefträger streichelte den Hund. Plötzlich biß er zu.
Wer beißt hier wen?

Tony Blair met President Yeltsin. The old man had just recovered from a heart attack.
Wer hatte einen Herzinfarkt?

Referenzresolution

Der Briefträger streichelte den Hund. Plötzlich biß er zu.
Wer beißt hier wen?

Tony Blair met President Yeltsin. The old man had just recovered from a heart attack.
Wer hatte einen Herzinfarkt?

Anwendungen:

- Informationsextraktion
- Question-Answering
- Summarisation
- Maschinelle Übersetzung
- ...

Wiederholung

- **Referenzausdrücke** (*die Queen*, *der Bus*, *eine Katze*, *er* . . .) referieren auf reale Entitäten
- Referenzausdrücke, die auf dieselbe Entität verweisen sind **koreferent**

Koreferenz und Anaphorik

- **Koreferenzkette** (coreference chain): eine Menge von Referenzausdrücken in einem Text, die koreferent sind
- **Anaphorik** (anaphora): ein Ausdruck verweist auf einen vorangegangenen Ausdruck (Antezedens)
- **Anapher** (anaphor): der zurückweisende Ausdruck (z.B. *sie*, *die Katze* . . . aber nicht *Peter*, *der Papst*, *the Queen* etc.)
- analog: Kataphorik (cataphora) für vorausweisende Ausdrücke
- **Koreferenz vs. Anaphorik**
  - cross-document coreference (=nicht anaphorisch)
  - Anaphern, die nicht koreferent sind (*Everybody has his own destiny.*)

Koreferenz und Anaphorik

- **Koreferenzkette** (coreference chain): eine Menge von Referenzausdrücken in einem Text, die koreferent sind
- **Anaphorik** (anaphora): ein Ausdruck verweist auf einen vorangegangenen Ausdruck (Antezedens)
- **Anapher** (anaphor): der zurückweisende Ausdruck (z.B. *sie*, *die Katze* . . . aber nicht *Peter*, *der Papst*, *the Queen* etc.)
- analog: Kataphorik (cataphora) für vorausweisende Ausdrücke
- **Koreferenz vs. Anaphorik**
  - cross-document coreference (=nicht anaphorisch)
  - Anaphern, die nicht koreferent sind (*Everybody has his own destiny.*)

Koreferenzresolution vs. Anaphernresolution

Koreferenzresolution: finde die Koreferenzketten in einem Text.

Anaphernresolution: finde das Antezedens einer Anapher.

Beispiel: Koreferenzresolution

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

Beispiel: Koreferenzresolution

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

Koreferenzketten (coreference chains):

Beispiel: Koreferenzresolution

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

Koreferenzketten (coreference chains):

- {Sophia Loren, she, the actress, her, she}

Beispiel: Koreferenzresolution

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

Koreferenzketten (coreference chains):

- {Sophia Loren, she, the actress, her, she}
- {Bono, the U2 singer }

Beispiel: Koreferenzresolution

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

Koreferenzketten (coreference chains):

- {Sophia Loren, she, the actress, her, she}
- {Bono, the U2 singer }
- {a thunderstorm}

Beispiel: Koreferenzresolution

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

Koreferenzketten (coreference chains):

- {Sophia Loren, she, the actress, her, she}
- {Bono, the U2 singer }
- {a thunderstorm}
- {a plane}

Beispiel: Anaphernresolution

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

Beispiel: Anaphernresolution

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

- *she* $\Rightarrow$ *Sophia Loren*

Beispiel: Anaphernresolution

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

- *she* $\Rightarrow$ *Sophia Loren*
- *the actress* $\Rightarrow$ *Sophia Loren*

Beispiel: Anaphernresolution

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

- *she* $\Rightarrow$ *Sophia Loren*
- *the actress* $\Rightarrow$ *Sophia Loren*
- *the U2 singer* $\Rightarrow$ *Bono*

Koreferenzresolution

Schwierigkeiten:

Koreferenzresolution

Schwierigkeiten:

- unterschiedliche Form $\not\Rightarrow$ verschiedene Referenten (*Sophia Loren* vs. *the actress* vs. *she*)

Koreferenzresolution

Schwierigkeiten:

- unterschiedliche Form $\not\Rightarrow$ verschiedene Referenten
  (*Sophia Loren* vs. *the actress* vs. *she*)
- gleiche Form $\not\Rightarrow$ gleiche Referenten
  (*die Katze*, *Michael Jackson* der Sänger vs. *Michael Jackson* der britische General)

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.

$\Rightarrow$ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.

$\Rightarrow$ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Peter told John that he is running the business for himself.

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.

$\Rightarrow$ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Peter told John that he is running the business for himself.

$\Rightarrow$ *syntactic constraints: himself* = *Peter*

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.
$\Rightarrow$ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Peter told John that he is running the business for himself.
$\Rightarrow$ *syntactic constraints: himself* = *Peter*

The cat did not come down from the tree. It was scared.

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.
$\Rightarrow$ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Peter told John that he is running the business for himself.
$\Rightarrow$ *syntactic constraints: himself* = *Peter*

The cat did not come down from the tree. It was scared.
$\Rightarrow$ selectional preferences: *it* = *the cat*

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.
$\Rightarrow$ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Peter told John that he is running the business for himself.
$\Rightarrow$ *syntactic constraints: himself* = *Peter*

The cat did not come down from the tree. It was scared.
$\Rightarrow$ selectional preferences: *it* = *the cat*

Jane told Mary she was in danger.

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.
⇒ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Peter told John that he is running the business for himself.
⇒ *syntactic constraints: himself* = *Peter*

The cat did not come down from the tree. It was scared.
⇒ selectional preferences: *it* = *the cat*

Jane told Mary she was in danger.
⇒ salience (Subjektposition): *she* = *Jane*

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.
$\Rightarrow$ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Peter told John that he is running the business for himself.
$\Rightarrow$ *syntactic constraints: himself* = *Peter*

The cat did not come down from the tree. It was scared.
$\Rightarrow$ selectional preferences: *it* = *the cat*

Jane told Mary she was in danger.
$\Rightarrow$ salience (Subjektposition): *she* = *Jane*

Jane told Mary SHE was in danger.

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.
⇒ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Peter told John that he is running the business for himself.
⇒ *syntactic constraints: himself* = *Peter*

The cat did not come down from the tree. It was scared.
⇒ selectional preferences: *it* = *the cat*

Jane told Mary she was in danger.
⇒ salience (Subjektposition): *she* = *Jane*

Jane told Mary SHE was in danger.
⇒ *Prosodie: she* = *Mary*

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.
⇒ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Peter told John that he is running the business for himself.
⇒ *syntactic constraints: himself* = *Peter*

The cat did not come down from the tree. It was scared.
⇒ selectional preferences: *it* = *the cat*

Jane told Mary she was in danger.
⇒ salience (Subjektposition): *she* = *Jane*

Jane told Mary SHE was in danger.
⇒ *Prosodie: she* = *Mary*

Jane warned Mary she was in danger.

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.
⇒ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Peter told John that he is running the business for himself.
⇒ *syntactic constraints: himself* = *Peter*

The cat did not come down from the tree. It was scared.
⇒ selectional preferences: *it* = *the cat*

Jane told Mary she was in danger.
⇒ salience (Subjektposition): *she* = *Jane*

Jane told Mary SHE was in danger.
⇒ *Prosodie: she* = *Mary*

Jane warned Mary she was in danger.
⇒ lexical semantics (*warned*): *she* = *Mary*

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.
⇒ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Peter told John that he is running the business for himself.
⇒ *syntactic constraints: himself* = *Peter*

The cat did not come down from the tree. It was scared.
⇒ selectional preferences: *it* = *the cat*

Jane told Mary she was in danger.
⇒ salience (Subjektposition): *she* = *Jane*

Jane told Mary SHE was in danger.
⇒ *Prosodie: she* = *Mary*

Jane warned Mary she was in danger.
⇒ lexical semantics (*warned*): *she* = *Mary*

Tony Blair met President Yeltsin. The old man had just recovered from a heart attack.

Ambiguität und desambiguierende Faktoren

Jane told Peter he was in danger.
⇒ Kongruenz (Genus, Numerus etc.): *he* = *Peter*

Peter told John that he is running the business for himself.
⇒ *syntactic constraints: himself* = *Peter*

The cat did not come down from the tree. It was scared.
⇒ selectional preferences: *it* = *the cat*

Jane told Mary she was in danger.
⇒ salience (Subjektposition): *she* = *Jane*

Jane told Mary SHE was in danger.
⇒ *Prosodie: she* = *Mary*

Jane warned Mary she was in danger.
⇒ lexical semantics (*warned*): *she* = *Mary*

Tony Blair met President Yeltsin. The old man had just recovered from a heart attack.
⇒ *Weltwissen: the old man* = *Yeltsin*

Schritte zur Anaphernresolution

1. identifiziere Anapher
   Schwierigkeiten: NPs, die keine Referenzausdrücke sind; pleonastisches *es* (*Es schneit.*) etc.
2. identifiziere potentielle Antezendenten
3. finde passendes Antezedens für jede Anapher

Schritte zur Anaphernresolution

1. identifiziere Anapher
   Schwierigkeiten: NPs, die keine Referenzausdrücke sind; pleonastisches *es* (*Es schneit.*) etc.
2. identifiziere potentielle Antezendenten
3. finde passendes Antezedens für jede Anapher

Ansätze zur Referenzresolution

Vor 1990 . . .

- Referenzresolution = Pronomenresolution
- regelbasiert (manuell erstellte Regeln)
- Beispiele:
  - SHRDLU (Winograd, 1972): komplexe Heuristiken (Fokus, Obliqueness etc.)
  - Hobbs's (1976, 1978): heuristisch gelenkte Suche in Syntaxbäumen
  - Centering-basiert (Brennan et al. 1987)

Nach 1990 . . .

- corpusbasiert (co-occurrence statistics, machine learning)
- auch Referenzresolution für nicht-pronominal Ausdrücke (definite NPs, bridging; z.B. Vieira & Poesio, 2000)

Wiederholung: Centering Theory

- **Backwards Looking Center**, C_b, verbindet U_n mit der vorangegangenen Äußerung U_{n-1}.
- **Forward Looking Centers**, C_f, bilden einen potientiellen Link mit der folgenden Äußerung U_{n+1}.
- Die partielle Ordnung der C_f wird u.A. durch die grammatische Role des Referenzausdrucks bestimmt.
- Das höchste Element im C_f einer Äußerung ist das präferierte Zentrum C_p.
- Das C_b einer Äußerung U_n ist das am höchsten gewertete Element des C_f in U_{n-1}, das in U_n realisiert ist.

Brennan et al. 1987

Zentrumstransitionen:

	$C_b(U_{n+1}) = C_b(U_n)$ or undefined $C_b(U_n)$	$C_b(U_{n+1}) \neq C_b(U_n)$
$C_b(U_{n+1}) = C_p(U_{n+1})$	Continue	Smooth-Shift
$C_b(U_{n+1}) \neq C_p(U_{n+1})$	Retain	Rough-Shift

Regeln:

Regel 1: wenn ein Element von $C_f(U_n)$ als Pronomen realisiert ist, dann muß auch $C_b(U_{n+1})$ pronominalisiert sein.

Regel 2: $Continue > Retain > Smooth - Shift > Rough - Shift$

Brennan et al. 1987

Anaphernresolution im Rahmen der Centering Theory:

1. generiere C_b, C_f für alle möglichen Referenzzuweisungen
2. Filter (Selektionspräferenzen, Regel 1, . . .)
3. ordne bei Regel 2 (d.h. löse Anapher so auf, daß die vorausgesagte Kohärenz des Textes maximiert wird)

Dagan & Itai (1990, 1991)

- corpus-basiert
- Referenzresolution auf Basis von Selektionspräferenzen
- Selektionspräferenzen automatisch aus Corpora extrahiert (co-occurrence statistics)

Dagan & Itai (1990, 1991)

They held tax money aside on the basis that the government said it was going to collect it.

Dagan & Itai (1990, 1991)

They held tax money aside on the basis that the government said it was going to collect it.

Dagan & Itai (1990, 1991)

They held tax money aside on the basis that the government said it was going to collect it.

Subjekt(it, collect)
Objekt(it, collect)

Dagan & Itai (1990, 1991)

They held tax money aside on the basis that the government said it was going to collect it.

Subjekt(it, collect)
Objekt(it, collect)

co-occurrence statistics:

Subjekt(money,collect) = 5
Subjekt(government,collect) = 198

Objekt(money,collect) = 149
Objekt(government,collect) = 0

Dagan & Itai (1990, 1991)

They held tax money aside on the basis that the government said it was going to collect it.

Subjekt(it, collect)
Objekt(it, collect)

co-occurrence statistics:

Subjekt(money,collect) = 5
Subjekt(government,collect) = 198

Objekt(money,collect) = 149
Objekt(government,collect) = 0

$\Rightarrow$ **it = government**
$\Rightarrow$ **it = money**

Dagan & Itai (1990, 1991)

Evaluation

- funktioniert nicht für alle Fälle
- aber kann gut mit einer allgemeineren Strategie verbunden werden (z.B. Hobbs's Baumsuche)

Lappin & Leass (1994)

- kein explizites semantisches oder pragmatisches Wissen
- mehrere Filter (syntaktisch, morphologisch etc.)
- saliency-based discourse model (wie stark ist ein potentieller Antzendens aktiviert?)
- corpus-basiertes Parametertuning

Saliency Faktoren (unterschiedlich gewichtet)

- wann erwähnt? (sentence recency)
- Subjektposition?
- Existenzkonstruktion? (*There are only a few restrictions . . .*)
- direktes Objekt?
- indirektes Objekt?
- eingebettete NP?
- . . .

Lappin & Leass (1994)

Ablauf (vereinfacht):

1. identifiziere mögliche Antezendenten
2. wende Filter an, um Antezendenten auszuschließen
3. ermittle saliency Werte für verbleibende Antezedenten
4. wähle Antezedens

Lappin & Leass (1994)

Ablauf (vereinfacht):

1. identifiziere mögliche Antezendenten
2. wende Filter an, um Antezendenten auszuschließen
3. ermittle saliency Werte für verbleibende Antezedenten
4. wähle Antezedens

Evaluation:

- 86% accuracy

Soon, Ng & Lim (1999)

- betreutes Maschinelles Lernen (Entscheidunsbäume)
- Trainingsdaten: annotierte Korpora, MUC (Message Understanding Conference)

Soon, Ng & Lim (1999)

- betreutes Maschinelles Lernen (Entscheidunsbäume)
- Trainingsdaten: annotierte Korpora, MUC (Message Understanding Conference)

Machine Learning Features:

- string similarity (*George W. Bush* vs. *Mr. Bush*)?
- Kongruenz (Numerus, Genus)?
- gleiche semantische (Wordnet-)Klasse (z.B. ORGANISATION, PERSON)?
- Pronomen, definite NP, demonstrative NP, Eigenname?

Soon, Ng & Lim (1999)

- betreutes Maschinelles Lernen (Entscheidunsbäume)
- Trainingsdaten: annotierte Korpora, MUC (Message Understanding Conference)

Machine Learning Features:

- string similarity (*George W. Bush* vs. *Mr. Bush*)?
- Kongruenz (Numerus, Genus)?
- gleiche semantische (Wordnet-)Klasse (z.B. ORGANISATION, PERSON)?
- Pronomen, definite NP, demonstrative NP, Eigenname?

Evaluation: 59% F-Score (vollautomatisches System, nicht nur Pronomen)

Demos

Lappin & Leass (1994):
`http://www-appn.comp.nus.edu.sg/%7Erpnlpir/cgi-bin/JavaRAP/JavaRAPdemo.html`

Mitkov (2002)
`http://clg.wlv.ac.uk/MARS/index.php`

Teil 2: Diskursinterpretation

Diskursparsing

Aufgabe:
identifiziere die Diskursstruktur eines Textes (Z.B. den RST Baum)

Anwendungen:

- information extraction
- question-answering
- text summarisation
- machine translation
- automatic essay scoring

Informationsquellen für Diskurstruktur

Diskurskonnektive (cue words) deuten Relationen an:
Peter kam zu spät, weil er den Bus verpaßt hatte.

Informationsquellen für Diskurstruktur

Diskurskonnektive (cue words) deuten Relationen an:
Peter kam zu spät, weil er den Bus verpaßt hatte.

Probleme:

Informationsquellen für Diskurstruktur

Diskurskonnektive (cue words) deuten Relationen an:
Peter kam zu spät, weil er den Bus verpaßt hatte.

Probleme:

- Diskursrelationen nicht immer durch Konnektive signalisiert

Informationsquellen für Diskurstruktur

Diskurskonnektive (cue words) deuten Relationen an:
Peter kam zu spät, weil er den Bus verpaßt hatte.

Probleme:

- Diskursrelationen nicht immer durch Konnektive signalisiert
  Peter kam zu spät. Er hatte den Bus verpaßt.

Informationsquellen für Diskurstruktur

Diskurskonnektive (cue words) deuten Relationen an:
Peter kam zu spät, weil er den Bus verpaßt hatte.

Probleme:

- Diskursrelationen nicht immer durch Konnektive signalisiert
  Peter kam zu spät. Er hatte den Bus verpaßt.
- *Konnektive zeigen nicht immer genau eine Relation an*

Informationsquellen für Diskurstruktur

Diskurskonnektive (cue words) deuten Relationen an:
Peter kam zu spät, weil er den Bus verpaßt hatte.

Probleme:

- Diskursrelationen nicht immer durch Konnektive signalisiert
  Peter kam zu spät. Er hatte den Bus verpaßt.
- *Konnektive zeigen nicht immer genau eine Relation an*
  It doesn't matter that it rains since we don't have time to go to the beach anyway. EXPLANATION
  She has worked in retail since she came to Britain. TEMPORAL

Informationsquellen für Diskurstruktur

Diskurskonnektive (cue words) deuten Relationen an:
Peter kam zu spät, weil er den Bus verpaßt hatte.

Probleme:

- Diskursrelationen nicht immer durch Konnektive signalisiert
  Peter kam zu spät. Er hatte den Bus verpaßt.
- *Konnektive zeigen nicht immer genau eine Relation an*
  It doesn't matter that it rains since we don't have time to go to the beach anyway. EXPLANATION
  She has worked in retail since she came to Britain. TEMPORAL
- *Ambiguität zwischen Diskurs- und Nicht-Diskursverwendung mancher Konnektive*

Informationsquellen für Diskurstruktur

Diskurskonnektive (cue words) deuten Relationen an:
Peter kam zu spät, weil er den Bus verpaßt hatte.

Probleme:

- Diskursrelationen nicht immer durch Konnektive signalisiert
  Peter kam zu spät. Er hatte den Bus verpaßt.
- *Konnektive zeigen nicht immer genau eine Relation an*
  It doesn't matter that it rains since we don't have time to go to the beach anyway. EXPLANATION
  She has worked in retail since she came to Britain. TEMPORAL

- *Ambiguität zwischen Diskurs- und Nicht-Diskursverwendung mancher Konnektive*
  Science has some definite conclusions about this. Yet, there are still many things we don't know. CONTRAST
  While there have been plans to extend the airport, nothing has been decided yet. NO DISCOURSE FUNCTION

Schritte zur Diskursstruktur

1. identifiziere die elementaren Diskurssegmente
2. konstruiere den Baum bottom-up
   - stelle fest, welche benachbarten Segment durch eine Relation verbunden sind
   - identifiziere die Relation, mit der die Segmente verbunden sind

Diskursparsing

- ursprünglich meist regelbasierte Systeme (manuell erstellte Regeln)
- seit späten 1990er: Corpora, die mit Diskursstruktur annotiert sind (RST Discourse Treebank; Penn Discourse Treebank; Potsdam Commentary Corpus)

  $\Rightarrow$ maschinelles Lernen von Diskursstruktur

Regelbasierte Ansätze

Wissensbasierte Modelle:

Hobbs et al. 1993, Asher & Lascarides 2003 etc.

- logik-basiert
- Weltwissen in Wissensbasis kodiert
- Diskursinterpretation ist eine Erweiterung von Satzinterpretation (d.h., Ziel ist es, die beste logische Form zu finden)

Wissensarme Systeme:

Marcu (1997), Polanyi et al. (2004), Corston-Oliver (1998), LeThan et al. (2004)

- Input: voll syntaktisch-analysierte Texte
- Heuristiken zur Erstellung der Diskursstruktur
- kein extensives semantisches Wissen (keine Wissensbasis)
- Oberflächenform (syntaktische Struktur, Deixis, Anaphern, cue words etc.) liefern Hinweise für die Diskursstruktur

Corpus-basierte Systeme

Marcu (1999), Baldridge & Lascarides (2005) etc.:

- betreutes maschinelles Lernen
- Trainingsdaten: z.B. RST Discourse Treebank
- Diskursparsing analog zum syntaktischen Parsing

Offene Probleme

. . . viele!

- wissensbasierte Systeme:
  - sehr viel Arbeit daher kaum implementiert (oder nur für eng umgrenzte Domänen)
  - im Großen vermutlich intractable
- wissensarme, heuristische Systeme:
  - relativ viel Arbeit
  - gut für relativ einfache Fälle, schlechte Abdeckung bei Relationen ohne Konnektive o.Ä.
- corpus-trainierte Systeme
  - wenig annotierte Trainingscorpora verfügbar
  - accuracies um die 60%

Ist vollständiges Diskursparsen für praktische Anwendungen überhaupt nötig?

Bibliographie

N. Asher and A. Lascarides

Logics of Conversation

Cambridge University Press, 2003.

J. Baldridge and A. Lascarides

Probabilistic Head-Driven Parsing for Discourse Structure.

CoNLL 2005, 2005.

S. Brennan and M. Friedman and C. Pollard

A centering approach to pronouns.

ACL 1987, 155-162, 1987.

S. Corston-Oliver

Identifying the linguistics correlates of rhetorical relations.

ACL Workshop in Discourse Relations and Discourse Markers, 1998.

J. Hobbs

Pronoun Resolution.

Research Report 76-1. New York: Department of Computer Science, City University of New York, 1976.

Bibliographie, cont'd

J. Hobbs

Resolving pronoun references.

Lingua, 44, 339–352, 1978.

J. Hobbs and M. Stickel and D. Appelt and P. Martin

Interpretation as Abduction.

Artificial Intelligence 63(1-2), 69–142, 1993.

S. Lappin and H. Leass

An algorithm for pronominal anaphora resolution.

Computational Linguistics, 20(4), 535–561, 1994.

H. Le Than and G. Abeysinghe and C. Huyck

Generating discourse structures for written text.

Coling-04, 2004.

D. Marcu

The Rhetorical Parsing, Summarization, and Generation of Natural Language Texts

PhD Thesis, Department of Computer Science, University of Toronto, 1997.

Bibliographie, cont'd

D. Marcu

A decision-based approach to rhetorical parsing.

ACL-99, 1999.

Ruslan Mitkov

Anaphora Resolution.

London: Longman, 2002.

L. Polanyi and C. Culy and M. van den Berg and G. L. Thione and D. Ahn

A rule-based approach to discourse parsing.

SIGDIAL Workshop in Discourse and Dialogue, 2004.

R. Vieira and M. Poesio

Processing definite descriptions in corpora.

In Botley, S. and McEnery, A. (eds) *Corpus-based and computational approaches to discourse anaphora*, 189–212, Amsterdam: John Benjamins, 2000.

Terry Winograd

Understanding Natural Language.

New York: Academic Press, 1972.