Coreference Resolution: Task, Linguistic Issues, Corpora, Evaluation

Michael Strube
Heidelberg Institute for Theoretical Studies gGmbH
Heidelberg, Germany
[Gaza] – [UPI]

The cheers and hails of happiness at a wedding in [Khan Younes in [the southern Gaza Strip]] turned into screams and moans of pain after [one of the celebrators] lost control of [[his] weapon, from which a number of bullets were released that killed [[the groom 's] brother] and hit three other relatives of [his], turning [the wedding] into a funeral in moments.

A medical source at [Nasser Hospital in [Khan Younes]] told [United Press International] that [Palestinian citizen Kamal Mohamed Al Bashiti –LRB– 33 years old –RRB–] died shortly after [he] was hit by a gunshot in the chest from [an accidental shooting at [[[his] brother 's] wedding in [Khan Younes]].

Three other relatives of [his] were also injured in [the accident], including two children.

[Gaza Strip] residents are suffering from the problem of aggravation from the weapons mess and security chaos resulting from the intense spread of weapons in the hands of security agency individuals and members of factions and families.

This is a common practice that claimed the lives of about 350 people last year, according to Palestinian human rights sources.

Mohamed Al Bashiti, 18 –LRB– years old –RRB–, a relative of [the victim], mentioned that while [the young men] were performing dances and popular dabke dances Thursday evening, and while [they] were in a state of intense rejoicing for [the wedding of [our relative, Majed Al Bashiti]], [one of the armed men] lost control of [[his] weapon, which [he] was trying to use to fire to celebrate and greet [the groom]].

The bullets went astray and hit a number of participants amid a state of panic and terror.

In moments, the wedding cheers turned into screams and the hitting of cheeks.

The participants of the party moved to [Nasser Hospital, which lay hundreds of meters away from the site of [the wedding]].

The tragic news came after a little while: the death of [[the groom 's] brother, who works in one of the security agencies].
Terminology

• Mentions (referring expressions): NL expression used by discourse participants to refer to entities
• Entities (referents): Entity that is referred to
• Coreference: Two or more mentions that are used to refer to the same entity are coreferent
• Anaphora: Reference to an entity that has been previously introduced in the discourse; the anaphor is used to refer to an entity already referred to by the antecedent
[Gaza] – [UPI]

The cheers and hails of happiness at a wedding in [Khan Younes in [the southern Gaza Strip]] turned into screams and moans of pain after [one of the celebrators] lost control of [[his] weapon , from which a number of bullets were released that killed [[the groom 's] brother] and hit three other relatives of [his] , turning [the wedding] into a funeral in moments].

A medical source at [Nasser Hospital in [Khan Younes]] told [United Press International] that [Palestinian citizen Kamal Mohamed Al Bashiti –LRB– 33 years old –RRB–] died shortly after [he] was hit by a gunshot in the chest from [an accidental shooting at [[his] brother 's] wedding in [Khan Younes]].

Three other relatives of [his] were also injured in [the accident] , including two children.

[Gaza Strip] residents are suffering from the problem of aggravation from the weapons mess and security chaos resulting from the intense spread of weapons in the hands of security agency individuals and members of factions and families.

This is a common practice that claimed the lives of about 350 people last year , according to Palestinian human rights sources.

Mohamed Al Bashiti , 18 –LRB– years old –RRB– , a relative of [the victim] , mentioned that while [the young men] were performing dances and popular dabke dances Thursday evening , and while [they] were in a state of intense rejoicing for [[the wedding of [our relative , Majed Al Bashiti]] , [one of the armed men] lost control of [[his] weapon , which [he] was trying to use to fire to celebrate and greet [the groom]].

The bullets went astray and hit a number of participants amid a state of panic and terror.

In moments , the wedding cheers turned into screams and the hitting of cheeks.

The participants of the party moved to [Nasser Hospital , which lay hundreds of meters away from the site of [the wedding]].

The tragic news came after a little while : the death of [[the groom 's] brother , who works in one of the security agencies].
Mention Types

• indefinite noun phrases
• definite noun phrases
• proper names
• pronouns
  • personal pronouns
  • possessive pronouns
  • reflexive pronouns
  • demonstrative pronouns

Note: I do not want to talk about plural pronouns with multiple singular antecedents, singular pronouns coreferent with a member of a set, generics, …
This is a common practice that T*-1 claimed the lives of about 350 people last year, according to Palestinian human rights sources.

Mohamed Al Bashiti, 18-LRB- years old -RRB-, a relative of [the victim], mentioned that while [the young men] were performing dances and popular dabke dances Thursday evening, and while [they] were in a state of intense rejoicing for [the wedding of [our relative, Majed Al Bashiti]], [one of the armed men] lost control of [[his] weapon, which [he] was trying *PRO*-1 to use *T*-2 *PRO*-3 to fire *PRO*-4 to celebrate and greet [the groom]]. The bullets went astray and hit a number of participants amid a state of panic and terror.

In moments, the wedding cheers turned into screams and the hitting of cheeks.

The participants of the party moved to [Nasser Hospital, which *T*-1 lay hundreds of meters away from the site of [the wedding]].

The tragic news *ICH*-1 came after a little while: the death of [[the groom 's] brother, who *T*-2 works in one of the security agencies].

[An elderly man] said, "[I] asked [them] *PRO*-1 not to [shoot] and asked [them] *PRO*-2 for God 's sake not to do [it], but [they] persisted in [it], until what *T*-3 happened happened.

A number of women suffered cases of fainting and shock after the tragic accident."

[Some Palestinians] are used to *PRO*-1 [shooting] in the air during [their] celebrations, especially weddings, as an expression of [their] happiness, in spite of [the repeated warnings 0 the Ministry of Interior has launched *T*-2 *PRO* to chase those who *T*-3 take part in [this practice]].

But [the warnings] were limited *-1 to announcements after each incident, without *PRO* translating into actual procedure in reality.

A researcher at the Palestinian Center for Human Rights, Yasser Abd Al Ghafoor, said this is not the first time 0 victims have fallen in similar incidents *T*-1, *PRO*-2 noting that the rights organizations observed several incidents in which victims fell and injuries occurred during wedding and celebration gunfire, one of the common misuses of weapons in the Gaza Strip *T*-3.

The city of Khan Younes had witnessed on the third of last July a similar incident, when [the girl, Amani Al Raqab], died by a bullet that *T*-1 hit [her] by mistake from a revolver 0 [[her] brother] was shooting *T*-2 off on [his] wedding day *T*-3.
Event Coreference

• Mentions can also be used to refer to events.
• Neuter pronoun *it*.
• Demonstrative pronouns have a preference for event antecedents.
• Not very frequent in text: less than 10% of the mentions have event antecedents.
• However, even common nouns can participate in event coreference.
In the Palestinian territories today, violence on violence again.

It began with the bombing of a school bus carrying children of Jewish settlers.

It ended with one of the most punishing Israeli responses in weeks.

Here's ABC's Gillian Findlay.

For two hours tonight, Israeli helicopters and ships fired missile after missile on Gaza, with 50 missiles in all, hitting police buildings, Palestinian security buildings, several offices of Yasser Arafat's political organization.

Unlike previous attacks, this one began without warning and early enough that people were still in the streets.

Gaza hospitals reported dozens of wounded, mostly bruises, doctors said, and shock.

Earlier today, it was Israelis in shock after a bomb exploded next to a bus filled with school children.

One of the Jewish settlements in Gaza.

The bus was armored, a usual precaution here, but the explosion was powerful.

Shrapnel killed two parents and injured several children seriously.

This road, a lifeline for several hundred Jewish settlers who live here, has been targeted dozens of times in recent weeks.

Today, the Palestinian Authority was quick to deny any responsibility, pointing out that the road is patrolled by Israeli soldiers.

The incident took place in an area which is under their control, under their security, jurisdiction, and they have to blame themselves.
[Gaza] – [UPI]

The cheers and hails of happiness at a wedding in [Khan Younes in [the southern Gaza Strip]] turned into screams and moans of pain after [one of the celebrators] lost control of [[his] weapon, from which a number of bullets were released that killed [[the groom 's] brother] and hit three other relatives of [his], turning [the wedding] into a funeral in moments.

A medical source at [Nasser Hospital in [Khan Younes]] told [United Press International] that [Palestinian citizen Kamal Mohamed Al Bashiti –LRB– 33 years old –RRB–] died shortly after [he] was hit by a gunshot in the chest from [an accidental shooting at [[[his] brother 's] wedding in [Khan Younes]].

Three other relatives of [his] were also injured in [the accident], including two children.

[Gaza Strip] residents are suffering from the problem of aggravation from the weapons mess and security chaos resulting from the intense spread of weapons in the hands of security agency individuals and members of factions and families.

This is a common practice that claimed the lives of about 350 people last year, according to Palestinian human rights sources.

Mohamed Al Bashiti, 18 –LRB– years old –RRB–, a relative of [the victim], mentioned that while [the young men] were performing dances and popular dabke dances Thursday evening, and while [they] were in a state of intense rejoicing for [the wedding of [our relative, Majed Al Bashiti]], [one of the armed men] lost control of [[his] weapon, which [he] was trying to use to fire to celebrate and greet [the groom]].

The bullets went astray and hit a number of participants amid a state of panic and terror.

In moments, the wedding cheers turned into screams and the hitting of cheeks.

The participants of the party moved to [Nasser Hospital, which lay hundreds of meters away from the site of [the wedding]].

The tragic news came after a little while: the death of [[the groom 's] brother, who works in one of the security agencies].
The cheers and hails of happiness at a wedding in [Khan Younes in [the southern Gaza Strip]] turned into screams and moans of pain after [one of the celebrators] lost control of [[his] weapon , from which a number of bullets were released that killed [[the groom's] brother] and hit three other relatives of [his] , turning [the wedding] into a funeral in moments .


Three other relatives of [his] were also injured in [the accident] , including two children .

[Gaza Strip] residents are suffering from the problem of aggravation from the weapons mess and security chaos resulting from the intense spread of weapons in the hands of security agency individuals and members of factions and families .

This is a common practice that claimed the lives of about 350 people last year , according to Palestinian human rights sources .

Mohamed Al Bashiti , 18 – LRB– years old –RRB– , a relative of [the victim] , mentioned that while [the young men] were performing dances and popular dabke dances Thursday evening , and while [they] were in a state of intense rejoicing for [the wedding of [our relative , Majed Al Bashiti]] , [one of the armed men] lost control of [[his] weapon , which [he] was trying to use to fire to celebrate and greet [the groom]] .

The bullets went astray and hit a number of participants amid a state of panic and terror .

In moments , the wedding cheers turned into screams and the hitting of cheeks .

The participants of the party moved to [Nasser Hospital] , which lay hundreds of meters away from the site of [the wedding] .

The tragic news came after a little while : the death of [[the groom's] brother , who works in one of the security agencies] .
The cheers and hails of happiness at a wedding in [Khan Younes in [the southern Gaza Strip]] turned into screams and moans of pain after [one of the celebrators] lost control of [[his] weapon, from which a number of bullets were released that killed [[the groom's] brother] and hit three other relatives of [his], turning [the wedding] into a funeral in moments.

A medical source at [Nasser Hospital in [Khan Younes]] told [United Press International] that [Palestinian citizen Kamal Mohamed Al Bashiti –LRB– 33 years old –RRB–] died shortly after [he] was hit by a gunshot in the chest from [an accidental shooting at [[his] brother's] wedding in [Khan Younes]].

Three other relatives of [his] were also injured in [the accident], including two children.

[Gaza Strip] residents are suffering from the problem of aggravation from the weapons mess and security chaos resulting from the intense spread of weapons in the hands of security agency individuals and members of factions and families.

This is a common practice that claimed the lives of about 350 people last year, according to Palestinian human rights sources.

Mohamed Al Bashiti, 18 –LRB– years old –RRB–, a relative of [the victim], mentioned that while [the young men] were performing dances and popular dabke dances Thursday evening, and while [they] were in a state of intense rejoicing for [the wedding of [our relative, Majed Al Bashiti]], [one of the armed men] lost control of [[his] weapon, which [he] was trying to use to fire to celebrate and greet [the groom]].

The bullets went astray and hit a number of participants amid a state of panic and terror.

In moments, the wedding cheers turned into screams and the hitting of cheeks.

The participants of the party moved to [Nasser Hospital], which lay hundreds of meters away from the site of [the wedding].

The tragic news came after a little while: the death of [[the groom's] brother, who works in one of the security agencies].
Domains

• News (MUC, ACE, OntoNotes):
  • highly ambiguous pronouns
  • many easily resolved proper names
  • very few hard cases which would require domain/world knowledge
  • OntoNotes contains some cases of event coreference
TEXT: [[I]] find no obvious source of [[her]] symptoms, especially [[this pain]].

Though [the defacography] is not normal, [[I]] do not think [that {findings}] are related to [[her]] symptoms.

[[These findings]] are consistent with normal variant.

[[I]] would not pursue {any type} of surgical repair or treatment and usually [we] treat [it] with Metamucil though [[the patient]] says that [[she]] has tried this, but with no success.

[[I]] have nothing further to offer [[this patient]] and explained this to [[her]] that [[she]] may be having [[this pain]] though [[I]] find {no source} for [[it]] and cannot offer any help for [[it]].

[[I]] did tell [[her]] that [[I]] would forward this information to [Dr. FFFFFFFFFFF;] AAAAAA Clinic – AAAAAAA; #### AAAAAA Lane, AAAAAAA, WI #### ###-####-####.

[[I]] offered to refer [[her]] to [one] of [[the gastroenterologists]] here at the BBBB Clinic, but [[she]] didn't think that [[they]] would have anything else to offer [[her]] that [[she]] hasn't had back home and declined any further referrals.

TEXT: [end section id="20112"]

TEXT: [start section id="20102"]

TEXT: Colon and Rectal Surgery

TEXT: Chief Resident of Dr. YYYYYYYYYYYYYYYYYYYYYY

TEXT: Rectal pain.

TEXT: [end section id="20102"]
This is a ##-year-old female sent for {intractable rectal pain.}

{She} has a rather complex history starting with {pelvic and abdominal pain} for quite some time and had no obvious diagnosis.

In November YYYY [[her]] [[doctors]] in MMMMMMMM, Louisiana, did an exploratory laparoscopy and saw only some [minimal endometriosis] [which] [[they]] didn't think was the cause of [[the pain.]]

However, [[she]] said that [[she]] was [pain-free] for two-weeks and then developed the fever and white count and [[the pain]] returned and [[she]] had [a repeat laparoscopy] [which] showed nothing.

{She} said that [[she]] continues to have low-grade fevers over the last three-months and has [a new-onset of a rectal pain] [which] starts shortly after [[she]] wakes up and builds up during {the day} and making [[her]] sick to [[her]] stomach and radiating over to the right-side.

{This type} of [rectal pain] is new since [[her]] [surgery] and [[it]] is there through {the day} and then when [[she]] goes to sleep, [[it]] goes away and the cycle continues again the next morning, shortly after [[she]] wakes up.

{She} moves [[her]] bowels and these are fairly solid.

{She} denies any blood in the stool.

{She} denies any masses or difficulty with defeation although [[she]] does have a history of constipation.

{She} had [[a colonoscopy]] and [[defecography]] done locally.
Relapsing-remitting multiple sclerosis

Catamenial migraines

I do not think that there is any doubt about the diagnosis here.

Dr. SSSSSSS has done an extremely thorough job in ruling out typical MS mimic conditions, and he has done the same workup we would have done here at the BBB Clinic.

I do not think that we can add anything to the diagnostic part of this illness.

As far as treatment, she is on a very appropriate medication, Avonex.

By now there were two studies comparing Avonex to higher dose preparations, namely, to Rebif and Betaseron.

Both studies have shown that higher dose interferon preparations are somewhat more effective in providing secondary prevention in relapsing-remitting MS.

One of these studies led to the eventual approval of Rebif in the United States.

With higher dose preparations, though, clearly the frequent injections do represent an inconvenience to the patient as well as the higher dose can result in more significant side-effects.

In the experience of Dr. Devon, Rebif actually is usually not associated with more side-effects than Avonex.

There is MRI evidence of inflammation on both of her MRIs in the last two years as well as she has a few T1 black hole formations which are thought to be associated with irreversible axonal loss.

Therefore, I think (for all) she should consider (if she can) with the help of available conservative treatment, which is Rebif in
[a few T1 black hole formations] [which] usually are thought to be associated with irreversible axonal loss.

Therefore, [[I]] think [[we]] should provide [[her]] with the best available preventive agent which is {Rebin} in [[our]] experience.

There is no urgency in making the switch and certainly one could follow an approach in which {we} see the next [MRI] scan in about 18 months and decide then whether [[she]] needs to be upgraded or not.

{However,} is [[she]] is willing to make that change earlier, that also seems appropriate.

[[I]] believe that no matter how [[she]] is doing clinically, even if things remain totally asymptomatic, [[she]] still should be followed with [[MRI]] scans in about every 18 months.

If that shows significantly increased [[disease]] activity, other treatment options could also be considered.

Switching to {Rebin} certainly does provide a broader spectrum prevention.

If that does not seem to be enough, occasionally mitoxantrone can be used as well which should not be considered in this case, since luckily [[her]] [MS] is not that severe.

There are several risk factors which are in [[her]] favor, including that [[she]] is [a woman] and [[she]] had mostly sensory exacerbations which all seem to suggest a better prognosis.

The age of onset is somewhat less favorable; however, overall, [[I]] am hoping that [[she]] is going to have a {relatively benign [disease]} course.

[[I]] let [[her]] know about the typical features of [the disease] including the immunological pathomechanism of [[relapsing-remitting MS]].

[We] also let [[her]] know about newer studies being conducted to promote nerve regeneration and myelin repair.

[[I]] provided some basic information about lifestyle modification issues and the usual management of fatigue.
Domains

• News (MUC, ACE, OntoNotes):
  • highly ambiguous pronouns
  • many easily resolved proper names
  • very few hard cases which would require domain/world knowledge
  • OntoNotes contains some cases of event coreference

• Medical Data (ODIE, i2b2):
  • first person singular pronouns almost always refer to physician
  • third person gendered singular pronouns almost always refer to patient
  • few proper names
  • quite a few very hard cases which require domain knowledge (which is readily available in the medical domain)
Annotated Corpora

• MUC-6, MUC-7 (Message Understanding Conference) – through LDC
• ACE 2002 - ACE 2005 (Automatic Content Extraction) – through LDC
• OntoNotes 4.0 (CoNLL Shared Task 2011), OntoNotes 5.0 (CoNNL Shared Task 2012 – Arabic, Chinese, English) – through LDC
• i2b2/VA 2011 (medical domain) – will be released later this year
• SemEval 2010 (Catalan, Dutch, English, German, Italian, Spanish) – through LDC
• Genia (documents about molecular biology) – http://www-tsujii.is.s.u-tokyo.ac.jp/GENIA/release/GENIA_MedCo_coreference_corpus_1.0.tar.gz
Annotation Tools

- MMAX2 (Müller & Strube, 2001); (Müller & Strube, 2006)
  http://mmax2.net
- PALinkA (Orăsan, 2003)
  http://clg.wlv.ac.uk/projects/PALinkA/
- GATE (Cunningham et al., 2002) http://gate.ac.uk/
- maybe also UIMA (Ferrucci & Lally, 2004)
  http://uima.apache.org/
Publically Available Systems

- BART (Poesio et al., 2007), (Versley et al., 2008)
  http://www.bart-coref.org/
- Reconcile (Stoyanov et al., 2009), (Stoyanov et al., 2010)
  http://www.cs.utah.edu/nlp/reconcile/
- Illinois coreference package (Bengtson & Roth, 2008), (Chang et al., 2011)
  http://cogcomp.cs.illinois.edu/page/software_view/18
- (Stanford coreference resolution system (Raghunathan et al., 2010), (Lee et al., 2011)
  http://nlp.stanford.edu/software/dcoref.shtml
Evaluation ...
... is Difficult

KEY

RESPONSE 1

RESPONSE 2

RESPONSE 3

RESPONSE 4
... is Difficult
... is Difficult

KEY
A  B  D
C  E
F  G

RESPONSE 1
A  B  D
C  E
F  G

RESPONSE 2
A  B  D
C  E
F  G

RESPONSE 3
A  B  D
C  E
F  G

RESPONSE 4
A  D
B
C  E
F  G
...is Difficult

because

• target concepts are sets/clusters
• some annotations include singletons, others don’t
• it is not quite clear what is better

hence, there is no single commonly agreed upon evaluation metric
Evaluation Metrics

• **MUC** (Vilain et al., 1995)
• **B-Cubed** (Bagga & Baldwin, 1998)
• **CEAF** (Luo, 2005)
• **BLANC** (Recasens & Hovy, 2010)
MUC Metric

• the oldest and most widely used metric
• focuses on links between (or, pairs of) mentions (the MUC data did not contain singletons)
• precision: \[
\frac{\#\text{common links in key and response}}{\#\text{links in response}}
\]
• recall: \[
\frac{\#\text{common links in key and response}}{\#\text{links in key}}
\]
• MUC metric prefers systems that have more mentions per entity (a system that creates a single entity with all mentions will get 100% recall without much degradation in precision)
• MUC metric ignores recall for singletons
B-Cubed Metric

- tries to address shortcomings of MUC metric
- focuses on mentions and computes precision and recall for each mention
  - $K$ is the key entity containing mention $M$
  - $R$ is the response entity containing mention $M$
  - precision for $M$: $\frac{|K \cap R|}{|R|}$
  - recall for $M$: $\frac{|K \cap R|}{|K|}$
- overall precision and recall are the average of the individual mention scores
CEAF Metric

• B-Cubed sometimes uses mentions multiple times
• CEAF metric addresses this shortcoming by aligning every response entity with at most one key entity
• finds the best one-to-one mapping between entities using an entity similarity metric
• hence, CEAF is an entity based measure
• depending on the similarity metric used there are two variants, entity based CEAF$_e$ and mention based CEAF$_m$

• precision: \[ \frac{\text{total similarity}}{\# \text{mentions in response}} \]
• recall: \[ \frac{\text{total similarity}}{\# \text{mentions in key}} \]
• BLANC uses a variation of the RAND index (metric for evaluating the quality of clustering algorithms)
CoNLL Metric

• there is no perfect metric
• compromise (already proposed by Denis & Baldridge (2009)): since MUC, B-Cubed and CEAF$_e$ take different perspectives, the unweighted mean among them may report fair results among competing systems
System Mentions

- until about two years ago, most systems received gold mentions as input
- in this case the proposed evaluation metrics including BLANC work fine
- however, if systems also have to recognize mentions, it is likely that key and response do not contain the same (number of) mentions
- evaluation metrics can be tricked by using a high recall mention tagger (Haghighi & Klein (2010) exploited this shortcoming – and received the Best Paper Award at NAACL’10 . . . )
System Mentions

• until about two years ago, most systems received gold mentions as input
• in this case the proposed evaluation metrics including BLANC work fine
• however, if systems also have to recognize mentions, it is likely that key and response do not contain the same (number of) mentions
• evaluation metrics can be tricked by using a high recall mention tagger (Haghighi & Klein (2010) exploited this shortcoming – and received the Best Paper Award at NAACL’10 . . .)
• Cai & Strube (2010) propose variants of B-Cubed and CEAF which take care of twinless mentions
• these variants were included into the CoNLL metric
A Few Notes on History

• Hobbs’ algorithm (Hobbs, 1978) – treewalk on syntactic structure
• Centering (Brennan et al., 1987; Strube, 1998; Tetreault, 2001)
• heuristics (Lappin & Leass, 1994)
• shallow, knowledge poor approaches (Baldwin, 1997; Mitkov, 1998)
A Few Notes on History

Source: Hobbs (1978)
A Few Notes on History

• Hobbs’ algorithm (Hobbs, 1978) – treewalk on syntactic structure
• Centering (Brennan et al., 1987; Strube, 1998; Tetreault, 2001)
• heuristics (Lappin & Leass, 1994)
• shallow, knowledge poor approaches (Baldwin, 1997; Mitkov, 1998)
A Few Notes on History

Salience factor types with initial weights

<table>
<thead>
<tr>
<th>Factor type</th>
<th>Initial weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence recency</td>
<td>100</td>
</tr>
<tr>
<td>Subject emphasis</td>
<td>80</td>
</tr>
<tr>
<td>Existential emphasis</td>
<td>70</td>
</tr>
<tr>
<td>Accusative emphasis</td>
<td>50</td>
</tr>
<tr>
<td>Indirect object and oblique complement emphasis</td>
<td>40</td>
</tr>
<tr>
<td>Head noun emphasis</td>
<td>80</td>
</tr>
<tr>
<td>Non-adverbial emphasis</td>
<td>50</td>
</tr>
</tbody>
</table>

Source: Lappin & Leass (1994)
A Few Notes on History

• Hobbs’ algorithm (Hobbs, 1978) – treewalk on syntactic structure
• Centering (Brennan et al., 1987; Strube, 1998; Tetreault, 2001)
• heuristics (Lappin & Leass, 1994)
• shallow, knowledge poor approaches (Baldwin, 1997; Mitkov, 1998)
A Few Notes on History

Baldwin (1997): high precision, rather low recall pronoun resolution system with the following strategy:

1. unique in discourse
2. reflexive
3. unique in current and prior sentences
4. possessive
5. unique in current sentence
6. unique subject
Homework

- report progress on the project; I expect you to have system output by now; evaluate it
- send this report to me (michael.strube@h-its.org) by May 31, 2012, 1pm
- (later today) link to the slides at http://michael.kimstrube.de/teaching.php
References

   Algorithms for scoring coreference chains.

Baldwin, Breck (1997).
   CogNIAC: High precision coreference with limited knowledge and linguistic resources.
   In *Proceedings of the ACL Workshop on Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Text, Madrid, Spain, July 1997*, pp. 38–45.

Bengtson, Eric & Dan Roth (2008).
   Understanding the value of features for coreference resolution.

   A centering approach to pronouns.

Cai, Jie & Michael Strube (2010).
   Evaluation metrics for end-to-end coreference resolution systems.

Chang, Kai-Wei, Rajhans Samdani, Alla Rozovskaya, Nick Rizzolo, Mark Sammons & Dan Roth (2011).
   Inference protocols for coreference resolution.

Cunningham, Hamish, Diana Maynard, Kalina Bontcheva & Valentin Tablan (2002).
   GATE: An architecture for development of robust HLT applications.

Denis, Pascal & Jason Baldridge (2009).
   Global joint models for coreference resolution and named entity classification.
UIMA: An architectural approach to unstructured information processing in the corporate research environment.  

Haghighi, Aria & Dan Klein (2010).  
Coreference resolution in a modular, entity centered model.  

Resolving pronominal references.  
Lingua, 44:311–338.

An algorithm for pronominal anaphora resolution.  

Stanford's multi-pass sieve coreference resolution system at the CoNLL-2011 shared task.  

Luo, Xiaoqiang (2005).  
On coreference resolution performance metrics.  

Robust pronoun resolution with limited knowledge.  

Müller, Christoph & Michael Strube (2001).  
MMAX: A tool for the annotation of multi-modal corpora.  

Müller, Christoph & Michael Strube (2006).  
Multi-level annotation of linguistic data with MMAX2.  
Orăsan, Constantin (2003).
PALinkA: A highly customizable tool for discourse annotation.

ELERFED: Final Report.

A multi-pass sieve for coreference resolution.

Recasens, Marta & Eduard Hovy (2010).
BLANC: Implementing the Rand index for coreference evaluation.

Stoyanov, Veselin, Claire Cardie, Nathan Gilbert, Ellen Riloff, David Buttler & David Hysom (2010).
Coreference resolution with Reconcile.

Conundrums in noun phrase coreference resolution: Making sense of the state-of-the-art.

Never look back: An alternative to centering.

A corpus-based evaluation of centering and pronoun resolution.

BART: A modular toolkit for coreference resolution.
In Companion Volume to the Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics, Columbus,