

Разрешение анафоры и кореферентности в текстах: методы, метрики и ресурсы

Максим Ионов
max.ionov@gmail.com

Goethe University Frankfurt / ОТиПЛ МГУ

Летняя школа АОТиАД, 27.07.2017

Outline

- 1 Понятия референции, кореферентности и анафоры
- 2 Постановка задачи и методы решения
 - Разрешение (местоименной) анафоры
 - Разрешение кореферентности
- 3 Метрики и ресурсы
 - Оценка качества: анафора

Референция

Референция — это способ указания на конкретные объекты действительности.

Референция

Референция — это способ указания на конкретные объекты действительности.

- Не обязательно конкретные:

(1) I saw *some people*.

Референция

Референция — это способ указания на конкретные объекты действительности.

- Не обязательно конкретные:
 - (1) I saw *some people*.
- Не обязательно действительности:
 - (2) I saw *a unicorn*.

Референция

Референция — это способ указания на конкретные объекты действительности.

- Не обязательно конкретные:
 - (1) I saw *some people*.
- Не обязательно действительности:
 - (2) I saw *a unicorn*.
- Не обязательно объекты:
 - (3) *The talk* is about reference.

Референция

Референция — это способ указания на конкретные объекты действительности.

- Не обязательно конкретные:
 - (1) I saw *some people*.
- Не обязательно действительности:
 - (2) I saw *a unicorn*.
- Не обязательно объекты:
 - (3) *The talk* is about reference.

Референция — это способ указывать на сущности в дискурсе.

Референция

- (4) I saw a unicorn in the forest. It was white. I hope I will see this animal again.

Референция

- (4) I saw a unicorn in the forest. It was white. I hope I will see this animal again.
- 3 сущности: *I, unicorn, forest*

Референция

(4) I saw a unicorn in the forest. It was white. I hope I will see this animal again.

- 3 сущности: *I, unicorn, forest*
- *a unicorn* — первое упоминание (discourse-new entity)

Референция

(4) I saw a unicorn in the forest. It was white. I hope I will see this animal again.

- 3 сущности: *I*, *unicorn*, *forest*
- *a unicorn* — первое упоминание (discourse-new entity)
- *the forest* указывает на конкретный лес

Референция

(4) I saw a unicorn in the forest. It was white. I hope I will see this animal again.

- 3 сущности: *I, unicorn, forest*
- *a unicorn* — первое упоминание (discourse-new entity)
- *the forest* указывает на конкретный лес
- *unicorn — it — this animal* указывает на одну и ту же сущность.

Референция

(4) I saw a unicorn in the forest. It was white. I hope I will see this animal again.

- 3 сущности: *I, unicorn, forest*
- *a unicorn* — первое упоминание (discourse-new entity)
- *the forest* указывает на конкретный лес
- *unicorn — it — this animal* указывает на одну и ту же сущность. Эти три именные группы **кореферентны** и образуют **кореферентную цепочку (coreferential chain)**

Референция

(4) I saw a unicorn in the forest. It was white. I hope I will see this animal again.

- 3 сущности: *I, unicorn, forest*
- *a unicorn* — первое упоминание (discourse-new entity)
- *the forest* указывает на конкретный лес
- *unicorn — it — this animal* указывает на одну и ту же сущность. Эти три именные группы **кореферентны** и образуют **кореферентную цепочку (coreferential chain)**
- местоимение, указывающее на другую именную группу — случай **анафоры**

Используемые понятия

- Упоминание — именная группа (с точки зрения дискурса)
- Референт — сущность, к которой относятся некоторые ИГ
- Кореферентность (кореферентная связь) — связь между двумя упоминаниями, которые относятся к одному и тому же элементу действительности

Используемые понятия

- Упоминание — именная группа (с точки зрения дискурса)
- Референт — сущность, к которой относятся некоторые ИГ
- Кореферентность (кореферентная связь) — связь между двумя упоминаниями, которые относятся к одному и тому же элементу действительности
- Первое упоминание — ИГ, вводящая новый элемент в дискурс
- Синглтон — ИГ, не вступающая в кореферентную связь

Используемые понятия

- Упоминание — именная группа (с точки зрения дискурса)
- Референт — сущность, к которой относятся некоторые ИГ
- Кореферентность (кореферентная связь) — связь между двумя упоминаниями, которые относятся к одному и тому же элементу действительности
- Первое упоминание — ИГ, вводящая новый элемент в дискурс
- Сингльтон — ИГ, не вступающая в кореферентную связь
- Анафор — местоимение, отсылающее к некоторой ИГ (второй элемент анафорической связи)
- Антецедент — первый элемент анафорической связи, ИГ, к которой относится местоимение

Исследования, связанные с референцией

- Философские аспекты

Исследования, связанные с референцией

- Философские аспекты
- Синтаксическая анафора

Исследования, связанные с референцией

- Философские аспекты
- Синтаксическая анафора
- Референциальный выбор

Исследования, связанные с референцией

- Философские аспекты
- Синтаксическая анафора
- Референциальный выбор
- (Дискурсивная) анафора и разрешение кореферентности

Исследования, связанные с референцией

- Философские аспекты
- Синтаксическая анафора
- Референциальный выбор
- (Дискурсивная) анафора и разрешение кореферентности
сегодняшний доклад

Синтаксическая анафора

В [Chomsky 1981] была представлена модификация генеративной грамматики — **теория управления и связывания (Government and Binding theory , GB)**. Значительная часть теории посвящена синтаксической анафоре — к чему может, а к чему не может относиться местоимение внутри предложения.

Синтаксическая анафора

В [Chomsky 1981] была представлена модификация генеративной грамматики — **теория управления и связывания (Government and Binding theory, GB)**. Значительная часть теории посвящена синтаксической анафоре — к чему может, а к чему не может относиться местоимение внутри предложения.

- (5)
- a. $Mary_i$ saw her_j .
 - b. * $Mary_i$ saw her_i .
 - c. * $Mary_i$ saw $herself_j$.
 - d. $Mary_i$ saw $herself_i$.
 - e. * $Mary_i$ saw $Mary_i$

Синтаксическая анафора

В [Chomsky 1981] была представлена модификация генеративной грамматики — **теория управления и связывания (Government and Binding theory, GB)**. Значительная часть теории посвящена синтаксической анафоре — к чему может, а к чему не может относиться местоимение внутри предложения.

- (5) a. $Mary_i$ saw her_j .
b. * $Mary_i$ saw her_i .
c. * $Mary_i$ saw $herself_j$.
d. $Mary_i$ saw $herself_i$.
e. * $Mary_i$ saw $Mary_i$

- Эти ограничения касаются только анафоры

Синтаксическая анафора

В [Chomsky 1981] была представлена модификация генеративной грамматики — **теория управления и связывания (Government and Binding theory, GB)**. Значительная часть теории посвящена синтаксической анафоре — к чему может, а к чему не может относиться местоимение внутри предложения.

- (5) a. $Mary_i$ saw her_j .
b. * $Mary_i$ saw her_i .
c. * $Mary_i$ saw $herself_j$.
d. $Mary_i$ saw $herself_i$.
e. * $Mary_i$ saw $Mary_i$

- Эти ограничения касаются только анафоры
- Только внутри предложения

Синтаксическая анафора

В [Chomsky 1981] была представлена модификация генеративной грамматики — **теория управления и связывания (Government and Binding theory, GB)**. Значительная часть теории посвящена синтаксической анафоре — к чему может, а к чему не может относиться местоимение внутри предложения.

- (5) a. $Mary_i$ saw her_j .
b. * $Mary_i$ saw her_i .
c. * $Mary_i$ saw $herself_j$.
d. $Mary_i$ saw $herself_i$.
e. * $Mary_i$ saw $Mary_i$

- Эти ограничения касаются только анафоры
 - Только внутри предложения
- ⇒ Не очень важно для обработки дискурса

Синтаксическая анафора

В [Chomsky 1981] была представлена модификация генеративной грамматики — **теория управления и связывания (Government and Binding theory, GB)**. Значительная часть теории посвящена синтаксической анафоре — к чему может, а к чему не может относиться местоимение внутри предложения.

- (5) a. $Mary_i$ saw her_j .
b. * $Mary_i$ saw her_i .
c. * $Mary_i$ saw $herself_j$.
d. $Mary_i$ saw $herself_i$.
e. * $Mary_i$ saw $Mary_i$

- Эти ограничения касаются только анафоры
 - Только внутри предложения
- ⇒ Не очень важно для обработки дискурса (но чрезвычайно важно для теории языка)

Референциальный выбор

В [Kibrik et al. 2010] (а также 2011 and 2013) представлена модель предсказания референциального выбора: выберет ли говорящий местоимение или полную ИГ в качестве упоминания. Некоторые важные признаки модели:

Референциальный выбор

В [Kibrik et al. 2010] (а также 2011 and 2013) представлена модель предсказания референциального выбора: выберет ли говорящий местоимение или полную ИГ в качестве упоминания. Некоторые важные признаки модели:

- Расстояние между анафором и антецедентом

Референциальный выбор

В [Kibrik et al. 2010] (а также 2011 and 2013) представлена модель предсказания референциального выбора: выберет ли говорящий местоимение или полную ИГ в качестве упоминания. Некоторые важные признаки модели:

- Расстояние между анафором и антецедентом
 - в предложениях
 - в абзацах
 - риторическое расстояние (в рамках теории риторических структур)

Референциальный выбор

В [Kibrik et al. 2010] (а также 2011 and 2013) представлена модель предсказания референциального выбора: выберет ли говорящий местоимение или полную ИГ в качестве упоминания. Некоторые важные признаки модели:

- Расстояние между анафором и антецедентом
 - в предложениях
 - в абзацах
 - риторическое расстояние (в рамках теории риторических структур)
- Протагонизм — центральность референта в дискурсе

Референциальный выбор

В [Kibrik et al. 2010] (а также 2011 and 2013) представлена модель предсказания референциального выбора: выберет ли говорящий местоимение или полную ИГ в качестве упоминания. Некоторые важные признаки модели:

- Расстояние между анафором и антецедентом
 - в предложениях
 - в абзацах
 - риторическое расстояние (в рамках теории риторических структур)
- Протагонизм — центральность референта в дискурсе
- Грамматические роли анафора и антецедента

Разрешение анафоры и кореферентности

Задача NLP по автоматическому установлению анафорических и кореферентных связей

Разрешение анафоры и кореферентности

Задача NLP по автоматическому установлению анафорических и кореферентных связей

- Важна для многих (почти всех) высокоуровневых задач NLP: извлечение фактов, машинный перевод, суммаризация, и т. д.

Разрешение анафоры и кореферентности

Задача NLP по автоматическому установлению анафорических и кореферентных связей

- Важна для многих (почти всех) высокоуровневых задач NLP: извлечение фактов, машинный перевод, суммаризация, и т. д.
- Первые исследования появились на заре NLP

Разрешение анафоры и кореферентности

Задача NLP по автоматическому установлению анафорических и кореферентных связей

- Важна для многих (почти всех) высокоуровневых задач NLP: извлечение фактов, машинный перевод, суммаризация, и т. д.
- Первые исследования появились на заре NLP
- Тем не менее, быстроразвивающаяся область, в особенности для low-resourced языков

Разрешение анафоры и кореферентности

Задача NLP по автоматическому установлению анафорических и кореферентных связей

- Важна для многих (почти всех) высокоуровневых задач NLP: извлечение фактов, машинный перевод, суммаризация, и т. д.
- Первые исследования появились на заре NLP
- Тем не менее, быстроразвивающаяся область, в особенности для low-resourced языков
- (как обычно) Первые алгоритмы были основаны на правилах, сейчас большинство основано на машинном обучении

Разрешение анафоры и кореферентности

Задача NLP по автоматическому установлению анафорических и кореферентных связей

- Важна для многих (почти всех) высокоуровневых задач NLP: извлечение фактов, машинный перевод, суммаризация, и т. д.
- Первые исследования появились на заре NLP
- Тем не менее, быстроразвивающаяся область, в особенности для low-resourced языков
- (как обычно) Первые алгоритмы были основаны на правилах, сейчас большинство основано на машинном обучении

Доклад посвящен именно этой задаче

Разрешение анафоры как частный случай разрешения кореферентности

- Анафора ← уровень текста, грамматики
- Кореферентность ← уровень дискурса

Разрешение анафоры как частный случай разрешения кореферентности

- Анафора ← уровень текста, грамматики
- Кореферентность ← уровень дискурса

С точки зрения NLP, задача разрешения анафоры — частный случай разрешения кореферентности.

Разрешение анафоры как частный случай разрешения кореферентности

- Анафора ← уровень текста, грамматики
- Кореферентность ← уровень дискурса

С точки зрения NLP, задача разрешения анафоры — частный случай разрешения кореферентности. (Poesio 2010):

*Interpreting anaphoric expressions is one of the most fundamental aspects of language interpretation. The study of **anaphora** and **anaphora resolution** (also known in Computational Linguistics as **coreference resolution**) has brought about*

Разрешение анафоры как частный случай разрешения кореферентности

Некоторые отличия:

Разрешение анафоры как частный случай разрешения кореферентности

Некоторые отличия:

- Методы оценки разрешения анафоры значительно проще

Разрешение анафоры как частный случай разрешения кореферентности

Некоторые отличия:

- Методы оценки разрешения анафоры значительно проще
- Анафору можно разрешать с меньшим количеством экстралингвистической информации

Разрешение анафоры как частный случай разрешения кореферентности

Некоторые отличия:

- Методы оценки разрешения анафоры значительно проще
- Анафору можно разрешать с меньшим количеством экстралингвистической информации
- Выделение всех упоминаний для разрешения анафоры тривиально

Разрешение анафоры как частный случай разрешения кореферентности

Некоторые отличия:

- Методы оценки разрешения анафоры значительно проще
- Анафору можно разрешать с меньшим количеством экстралингвистической информации
- Выделение всех упоминаний для разрешения анафоры тривиально (более или менее)

Разрешение анафоры как частный случай разрешения кореферентности

Некоторые отличия:

- Методы оценки разрешения анафоры значительно проще
- Анафору можно разрешать с меньшим количеством экстралингвистической информации
- Выделение всех упоминаний для разрешения анафоры тривиально (более или менее)

(6) It is raining.

Задача

3. Великолепная «Школа злословия» вернулась в эфир после летних каникул в новом формате.
4. В истории программы это уже не первый «ребрендинг».
5. Сейчас с трудом можно припомнить, что начиналась «Школа...» на канале «Культура» как стандартное ток-шоу, которое отличалось от других «кухонными» обсуждениями гостя, что называется – «за глаза», и неожиданными персонами в качестве ведущих.
6. Писательница Татьяна Толстая и сценаристка Дуня Смирнова вроде бы не вполне соответствовали принятым на российском телевидении стандартам телеведущих.
7. Впрочем, на канале «Культура» в роли телеведущих выступают и писатели, и композиторы, так что в этом ничего сверхъестественного не было, а идея кухонных обсуждений не слишком прижилась, и некоторые выпуски программы обходились практически без них.

Задача

- 3: Великолепная «Школа злословия» вернулась в эфир после летних каникул в новом формате.
- 4: В истории программы это уже не первый «ребрендинг».
- 5: Сейчас с трудом можно припомнить, что начиналась «Школа...» на канале «Культура» как стандартное ток-шоу, которое отличалось от других «кухонными» обсуждениями гостя, что называется – «за глаза», и неожиданными персонами в качестве ведущих.
- 6: Писательница Татьяна Голстая и сценаристка Дуня Смирнова вроде бы не вполне соответствовали принятым на российском телевидении стандартам телеведущих.
- 7: Впрочем, на канале «Культура» в роли телеведущих выступают и писатели, и композиторы, так что в этом ничего сверхъестественного не было, а идея кухонных обсуждений не слишком прижилась, и некоторые выпуски программы обходились практически без них.

Outline

- 1 Понятия референции, кореферентности и анафоры
- 2 **Постановка задачи и методы решения**
 - Разрешение (местоименной) анафоры
 - Разрешение кореферентности
- 3 Метрики и ресурсы
 - Оценка качества: анафора

Contents

- 1 Понятия референции, кореферентности и анафоры
- 2 **Постановка задачи и методы решения**
 - Разрешение (местоименной) анафоры
 - Разрешение кореферентности
- 3 Метрики и ресурсы
 - Оценка качества: анафора

Разрешение анафоры: правилый подход

- Формулировка задачи: для каждого (?) местоимения в тексте найти его antecedent.

Разрешение анафоры: правилый подход

- Формулировка задачи: для каждого (?) местоимения в тексте найти его антецедент.
- Правильовые подходы: модель *generate-filter-rank*:

Разрешение анафоры: правилый подход

- Формулировка задачи: для каждого (?) местоимения в тексте найти его antecedent.
- Правилые подходы: модель *generate-filter-rank*:
 - Для каждого анафора — выделение ИГ-кандидатов

Разрешение анафоры: правилый подход

- Формулировка задачи: для каждого (?) местоимения в тексте найти его antecedent.
- Правилые подходы: модель *generate-filter-rank*:
 - Для каждого анафора — выделение ИГ-кандидатов
 - Фильтрация кандидатов, нарушающих ограничения (например, согласование по роду и числу)

Разрешение анафоры: правилый подход

- Формулировка задачи: для каждого (?) местоимения в тексте найти его antecedent.
- Правилые подходы: модель *generate-filter-rank*:
 - Для каждого анафора — выделение ИГ-кандидатов
 - Фильтрация кандидатов, нарушающих ограничения (например, согласование по роду и числу)
 - Ранжирование оставшихся гипотез согласно предпочтениям (активированность референта, расстояние между анафором и antecedentом)

Разрешение анафоры: правилый подход

- Формулировка задачи: для каждого (?) местоимения в тексте найти его antecedent.
- Правилые подходы: модель *generate-filter-rank*:
 - Для каждого анафора — выделение ИГ-кандидатов
 - Фильтрация кандидатов, нарушающих ограничения (например, согласование по роду и числу)
 - Ранжирование оставшихся гипотез согласно предпочтениям (активированность референта, расстояние между анафором и antecedentом)
- Отбор кандидатов стоит ограничивать по расстоянию: > 90% antecedentов находятся в том же или предыдущем предложении, что и анафор (напр., [Hobbs 1978]).

Разрешение анафоры: правилый подход

Результаты для простых эвристик (на материале корпуса RuCor):

- Бейзлайн: выбор ближайшего ИГ
- Более сложный бейзлайн: выбор ближайшего совпадающего по роду и числу ИГ + простые эвристики для разных типов местоимений

Table 1: Качество разрешения анафоры при подходе, основанном на правилах

	P	R	F_1
Baseline	37.2%	35.7%	36.4%
Baseline + rules	56.5%	51.6%	53.9%

Разрешение анафоры: использование машинного обучения

- Бинарная классификация для каждой гипотезы
антецедент—анафор

Разрешение анафоры: использование машинного обучения

- Бинарная классификация для каждой гипотезы
антецедент—анафор
- Можно применять предварительную фильтрацию,
используя строгие ограничения

Разрешение анафоры: использование машинного обучения

- Бинарная классификация для каждой гипотезы
антецедент—анафор
- Можно применять предварительную фильтрацию,
используя строгие ограничения
- Проблема несбалансированного множества: отрицательных
примеров значительно больше

Разрешение анафоры: использование машинного обучения

Признаки для классификации:

Разрешение анафоры: использование машинного обучения

Признаки для классификации:

- LENDIST: признаки длины и расстояния.
 - Длина ИГ-кандидата в символах.
 - Длина кандидата в словах.
 - Расстояние между местоимением и ИГ в символах.
 - Расстояние между местоимением и ИГ в словах.
 - Расстояние между местоимением и ИГ в группах.

Разрешение анафоры: использование машинного обучения

Признаки для классификации:

- AGR: признаки согласования между ИГ и местоимением.
 - Показатель числа ИГ-кандидата.
 - Показатель числа местоимения.
 - Совпадают ли показатели числа у кандидата и местоимения.

Разрешение анафоры: использование машинного обучения

Признаки для классификации:

- SALIENCE: признаки, коррелирующие с салиентностью (активированностью).
 - Падеж ИГ.
 - Падеж местоимения.
 - Совпадают ли падежи у ИГ и местоимения.
 - Является ли кандидат именем собственным
 - Количество вхождений кандидата в тексте.
 - Тип местоимения.
 - Само местоимение.

Contents

- 1 Понятия референции, кореферентности и анафоры
- 2 **Постановка задачи и методы решения**
 - Разрешение (местоименной) анафоры
 - Разрешение кореферентности
- 3 Метрики и ресурсы
 - Оценка качества: анафора

Разрешение кореферентности: постановка задачи

- Формулировка задачи: разбить множество упоминаний на кореферентные цепочки.

Разрешение кореферентности: постановка задачи

- Формулировка задачи: разбить множество упоминаний на кореферентные цепочки.
- Получение множества упоминаний:

Разрешение кореферентности: постановка задачи

- Формулировка задачи: разбить множество упоминаний на кореферентные цепочки.
- Получение множества упоминаний:
 - Все ИГ: сильно зависит от морфологического анализа и более «нижних» уровней анализа

Разрешение кореферентности: постановка задачи

- Формулировка задачи: разбить множество упоминаний на кореферентные цепочки.
- Получение множества упоминаний:
 - Все ИГ: сильно зависит от морфологического анализа и более «нижних» уровней анализа
 - ИГ, данные извне: не подходит для практического применения, но позволяет тестировать непосредственно алгоритм разрешения кореферентности

Разрешение кореферентности: постановка задачи

- Формулировка задачи: разбить множество упоминаний на кореферентные цепочки.
- Получение множества упоминаний:
 - Все ИГ: сильно зависит от морфологического анализа и более «нижних» уровней анализа
 - ИГ, данные извне: не подходит для практического применения, но позволяет тестировать непосредственно алгоритм разрешения кореферентности
 - Предварительно отфильтрованные (кореферентные) ИГ: без синглтонов

Разрешение кореферентности: различные модели

- MENTION-PAIR:
 - Генерируются пары-гипотезы
 - Для каждой пары решается, кореферентны ли ИГ
 - Кореферентные пары группируются в цепочки

Разрешение кореферентности: различные модели

- MENTION-PAIR:
 - Генерируются пары-гипотезы
 - Для каждой пары решается, кореферентны ли ИГ
 - Кореферентные пары группируются в цепочки
- ENTITY-MENTION:
 - Генерируется список ИГ
 - Для каждой ИГ решается, к какой цепочке ее отнести (либо создать новую)

Разрешение кореферентности: различные модели

- MENTION-PAIR:
 - Генерируются пары-гипотезы
 - Для каждой пары решается, кореферентны ли ИГ
 - Кореферентные пары группируются в цепочки
- ENTITY-MENTION:
 - Генерируется список ИГ
 - Для каждой ИГ решается, к какой цепочке ее отнести (либо создать новую)
- Есть и другие модели: глобальное обучение, турнирная модель, etc.

Разрешение кореферентности: MENTION-PAIR

- Простейшая модель для разрешения кореферентности с использованием ML

Разрешение кореферентности: MENTION-PAIR

- Простейшая модель для разрешения кореферентности с использованием ML
- Генерация гипотез: для каждой пары анафор + антецедент из обучающей выборки:

Разрешение кореферентности: MENTION-PAIR

- Простейшая модель для разрешения кореферентности с использованием ML
- Генерация гипотез: для каждой пары анафор + антецедент из обучающей выборки:
 - Положительный пример: сама пара
 - Отрицательные примеры: анафор + ИГ, расположенные между положительной парой

Разрешение кореферентности: MENTION-PAIR

- Простейшая модель для разрешения кореферентности с использованием ML
- Генерация гипотез: для каждой пары анафор + antecedent из обучающей выборки:
 - Положительный пример: сама пара
 - Отрицательные примеры: анафор + ИГ, расположенные между положительной парой
- Объединение пар в цепочки:
 - CLOSE-FIRST: в цепочку объединяется анафор и ближайший возможный antecedent
 - BEST-FIRST: в цепочку объединяется пара с наибольшей вероятностью

Разрешение кореферентности: MENTION-PAIR

- Простейшая модель для разрешения кореферентности с использованием ML
- Генерация гипотез: для каждой пары анафор + антецедент из обучающей выборки:
 - Положительный пример: сама пара
 - Отрицательные примеры: анафор + ИГ, расположенные между положительной парой
- Объединение пар в цепочки:
 - CLOSE-FIRST: в цепочку объединяется анафор и ближайший возможный антецедент
 - BEST-FIRST: в цепочку объединяется пара с наибольшей вероятностью
 - и многие другие стратегии

Разрешение кореферентности: MENTION-PAIR и ENTITY-MENTION

- Модель имеет серьезный недостаток: она не использует глобальную информацию, что приводит к «глупым» ошибкам: пары *Hilary Clinton* → *Clinton* и *Clinton* → *he* могут объединиться в одну цепочку
- Модель ENTITY-MENTION является наиболее приближенной к когнитивным процессам

Разрешение кореферентности: признаки

- 1 Строковые: полное или частичное совпадение ИГ, совпадение вершин, альтернативное написание, расстояние между ИГ.

Разрешение кореферентности: признаки

- 1 Строковые: полное или частичное совпадение ИГ, совпадение вершин, альтернативное написание, расстояние между ИГ.
- 2 Грамматические признаки ИГ-кандидата или анафора: является ли ИГ местоимением, определенность, наличие указательных местоимений.

Разрешение кореферентности: признаки

- 1 Строковые: полное или частичное совпадение ИГ, совпадение вершин, альтернативное написание, расстояние между ИГ.
- 2 Грамматические признаки ИГ-кандидата или анафора: является ли ИГ местоимением, определенность, наличие указательных местоимений.
- 3 Грамматические признаки сочетания ИГ-кандидата и анафора: согласование по лицу, числу, являются ли обе ИГ именами собственными, находится ли одна ИГ в аппозитивной или предикативной позиции относительно другой, синтаксические роли ИГ.

Разрешение кореферентности: признаки

- 1 Строковые: полное или частичное совпадение ИГ, совпадение вершин, альтернативное написание, расстояние между ИГ.
- 2 Грамматические признаки ИГ-кандидата или анафора: является ли ИГ местоимением, определенность, наличие указательных местоимений.
- 3 Грамматические признаки сочетания ИГ-кандидата и анафора: согласование по лицу, числу, являются ли обе ИГ именами собственными, находится ли одна ИГ в аппозитивной или предикативной позиции относительно другой, синтаксические роли ИГ.
- 4 Семантические свойства: семантические отношения двух ИГ (синонимия, гиперонимия), совпадают ли семантические классы двух ИГ.

Разрешение кореферентности: признаки

- 1 Строковые: полное или частичное совпадение ИГ, совпадение вершин, альтернативное написание, расстояние между ИГ.
- 2 Грамматические признаки ИГ-кандидата или анафора: является ли ИГ местоимением, определенность, наличие указательных местоимений.
- 3 Грамматические признаки сочетания ИГ-кандидата и анафора: согласование по лицу, числу, являются ли обе ИГ именами собственными, находится ли одна ИГ в аппозитивной или предикативной позиции относительно другой, синтаксические роли ИГ.
- 4 Семантические свойства: семантические отношения двух ИГ (синонимия, гиперонимия), совпадают ли семантические классы двух ИГ.
- 5 Типы упоминаний: является ли ИГ синглтоном, является

Разрешение кореферентности: признаки

Outline

- 1 Понятия референции, кореферентности и анафоры
- 2 Постановка задачи и методы решения
 - Разрешение (местоименной) анафоры
 - Разрешение кореферентности
- 3 Метрики и ресурсы
 - Оценка качества: анафора

Contents

- 1 Понятия референции, кореферентности и анафоры
- 2 Постановка задачи и методы решения
 - Разрешение (местоименной) анафоры
 - Разрешение кореферентности
- 3 Метрики и ресурсы
 - Оценка качества: анафора

Оценка качества: анафора

- Anaphora resolution task is simple: for every pronoun restore its antecedent

Оценка качества: анафора

- Anaphora resolution task is simple: for every pronoun restore its antecedent
- Classification task: classify all pronouns as a match for the antecedent

Оценка качества: анафора

- Anaphora resolution task is simple: for every pronoun restore its antecedent
- Classification task: classify all pronouns as a match for the antecedent
- We can simply use the Accuracy metric (% of correctly resolved anaphoric links)...

Оценка качества: анафора

- Anaphora resolution task is simple: for every pronoun restore its antecedent
- Classification task: classify all pronouns as a match for the antecedent
- We can simply use the Accuracy metric (% of correctly resolved anaphoric links) . . .
- But we want to know the recall as well

Оценка качества: анафора

- Anaphora resolution task is simple: for every pronoun restore its antecedent
- Classification task: classify all pronouns as a match for the antecedent
- We can simply use the Accuracy metric (% of correctly resolved anaphoric links) . . .
- But we want to know the recall as well
- Some algorithms may resolve *something* for each pronoun

Оценка качества: анафора

- Anaphora resolution task is simple: for every pronoun restore its antecedent
- Classification task: classify all pronouns as a match for the antecedent
- We can simply use the Accuracy metric (% of correctly resolved anaphoric links) . . .
- But we want to know the recall as well
- Some algorithms may resolve *something* for each pronoun
- Some may yield only precise results

Оценка качества: анафора

- Anaphora resolution task is simple: for every pronoun restore its antecedent
 - Classification task: classify all pronouns as a match for the antecedent
 - We can simply use the Accuracy metric (% of correctly resolved anaphoric links) . . .
 - But we want to know the recall as well
 - Some algorithms may resolve *something* for each pronoun
 - Some may yield only precise results
- ⇒ A combination of precision and recall gives more transparent results

Evaluation metrics: anaphora

- TP # of pairs correctly classified as true
- FP # of pairs incorrectly classified as true
- FN # of pairs incorrectly classified as false

Evaluation metrics: anaphora

- TP # of pairs correctly classified as true
- FP # of pairs incorrectly classified as true
- FN # of pairs incorrectly classified as false
- From this we can get precision, recall and F-measures:

$$P = \frac{N_{correct}}{N_{found}}$$

$$R = \frac{N_{correct}}{N_{total}}$$

$$F_1 = \frac{2 \cdot P \cdot R}{P + R}$$

Evaluation metrics: coreference

Coreference resolution task can't be evaluated this way, since chains are sets:

Evaluation metrics: coreference

Coreference resolution task can't be evaluated this way, since chains are sets:

- Overlap

Evaluation metrics: coreference

Coreference resolution task can't be evaluated this way, since chains are sets:

- Overlap
- Two chains in GS, one chain in test results

Evaluation metrics: coreference

Coreference resolution task can't be evaluated this way, since chains are sets:

- Overlap
- Two chains in GS, one chain in test results
- One chain in GS, two chains in test results

Evaluation metrics: coreference

Coreference resolution task can't be evaluated this way, since chains are sets:

- Overlap
- Two chains in GS, one chain in test results
- One chain in GS, two chains in test results
- ...

Evaluation metrics: coreference

There are various ways to evaluate this:

Evaluation metrics: coreference

There are various ways to evaluate this:

- Jaccard coefficient

Evaluation metrics: coreference

There are various ways to evaluate this:

- Jaccard coefficient
- MUC-score

Evaluation metrics: coreference

There are various ways to evaluate this:

- Jaccard coefficient
- MUC-score
- B³

Evaluation metrics: coreference

$$Jac(C_1, C_2) = \frac{C_1 \cap C_2}{C_1 \cup C_2},$$

where C_1 and C_2 are chains represented by sets

MUC-score

- All mentions forms equivalency classes (EC)

MUC-score

- All mentions forms equivalency classes (EC)
- Each class represents one coreferential chain

MUC-score

- All mentions forms equivalency classes (EC)
- Each class represents one coreferential chain
- Metric value means the differences in the number of connections: the minimum number of links between mentions to be inserted or deleted when mapping a system response to a gold standard key set.

MUC-score

S — Set of EC in GS

R — Set of EC in test results

$p(S)$ — partition of S relatively to the result

$c(S) = (|S| - 1)$ — minimal number of links to get S

$m(S) = (|p(S)| - 1)$ — number of missing links in R compared with GS

MUC-score

S — Set of EC in GS

R — Set of EC in test results

$p(S)$ — partition of S relatively to the result

$c(S) = (|S| - 1)$ — minimal number of links to get S

$m(S) = (|p(S)| - 1)$ — number of missing links in R compared with GS

$$P = \frac{\sum_i (|R_i| - |p(R_i)|)}{\sum_i (|R_i| - 1)}$$

$$R = \frac{\sum_i (|S_i| - |p(S_i)|)}{\sum_i (|S_i| - 1)}$$

B-CUBED score

MUC-score is not perfect:

B-CUBED score

MUC-score is not perfect:

- Singletons (chains of length 1) are not accounted at all (there are no connections)

B-CUBED score

MUC-score is not perfect:

- Singletons (chains of length 1) are not accounted at all (there are no connections)
- All mistakes are equal: erroneous adding one element to a big chain is equal to merging two big chains

B-CUBED score

MUC-score is not perfect:

- Singletons (chains of length 1) are not accounted at all (there are no connections)
- All mistakes are equal: erroneous adding one element to a big chain is equal to merging two big chains

B^3 score was created to overcome this:

B-CUBED score

MUC-score is not perfect:

- Singletons (chains of length 1) are not accounted at all (there are no connections)
- All mistakes are equal: erroneous adding one element to a big chain is equal to merging two big chains

B³ score was created to overcome this:

- Count precision and recall separately for every referent

B-CUBED score

MUC-score is not perfect:

- Singletons (chains of length 1) are not accounted at all (there are no connections)
- All mistakes are equal: erroneous adding one element to a big chain is equal to merging two big chains

B^3 score was created to overcome this:

- Count precision and recall separately for every referent
- Result is a weighted sums of those scores

B-CUBED score

$$P_i = \frac{N \text{ true elements in test results with a referent}_i}{N \text{ elements in a chain with referent}_i}$$
$$R_i = \frac{N \text{ true elements in test results with a referent}_i}{N \text{ elements in a GS chain with referent}_i}$$

B-CUBED score

$$P_i = \frac{N \text{ true elements in test results with a referent}_i}{N \text{ elements in a chain with referent}_i}$$

$$R_i = \frac{N \text{ true elements in test results with a referent}_i}{N \text{ elements in a GS chain with referent}_i}$$

Overall quality:

$$P = \sum_{i=1}^N w_i \cdot P_i$$

$$R = \sum_{i=1}^N w_i \cdot R_i$$

B-CUBED score

$$P_i = \frac{N \text{ true elements in test results with a referent}_i}{N \text{ elements in a chain with referent}_i}$$

$$R_i = \frac{N \text{ true elements in test results with a referent}_i}{N \text{ elements in a GS chain with referent}_i}$$

Overall quality:

$$P = \sum_{i=1}^N w_i \cdot P_i$$

$$R = \sum_{i=1}^N w_i \cdot R_i$$

Weights can be set depending on the task

- B^3 can count some chains more than once, and this makes the results less fair

CEAF

- B^3 can count some chains more than once, and this makes the results less fair
- Constrained Entity-Alignment F-Measure (CEAF)

CEAF

- B^3 can count some chains more than once, and this makes the results less fair
- Constrained Entity-Alignment F-Measure (CEAF)
- A lot of varieties: $CEAF_m$, $CEAF_e$, etc.

CEAF

- B^3 can count some chains more than once, and this makes the results less fair
- Constrained Entity-Alignment F-Measure (CEAF)
- A lot of varieties: $CEAF_m$, $CEAF_e$, etc.
- Compares the sets of aligned coreference chains

Comparison of the metrics

- MUC penalizes missing links:

Comparison of the metrics

- MUC penalizes missing links: A simple baseline, all mentions form one chain, gives us MUC F_1 -score $\approx 85\%$

Comparison of the metrics

- MUC penalizes missing links: A simple baseline, all mentions form one chain, gives us MUC F_1 -score $\approx 85\%$
- B^3 measures quality for an *average chain*

Comparison of the metrics

- MUC penalizes missing links: A simple baseline, all mentions form one chain, gives us MUC F_1 -score $\approx 85\%$
- B^3 measures quality for an *average chain*
- CEAF measures the *similarity* between two sets of chains

Comparison of the metrics

- CoNLL reference scorer

Comparison of the metrics

- CoNLL reference scorer
- <https://github.com/conll/reference-coreference-scorers>

Comparison of the metrics

- CoNLL reference scorer
- `https://github.com/conll/reference-coreference-scorers`
- A reliable way to compute all the scores

Corpora

- There are a lot of corpora for a number of languages:

Corpora

- There are a lot of corpora for a number of languages:
 - English: MUC, ACE, ...

Corpora

- There are a lot of corpora for a number of languages:
 - English: MUC, ACE, ...
 - German: TüBa-D/Z, ...

Corpora

- There are a lot of corpora for a number of languages:
 - English: MUC, ACE, ...
 - German: TüBa-D/Z, ...
 - and many others
- The only available coreference corpus is RuCoref (Toldova et al. 2014)

Corpora

- There are a lot of corpora for a number of languages:
 - English: MUC, ACE, ...
 - German: TüBa-D/Z, ...
 - and many others
- The only available coreference corpus is RuCoref (Toldova et al. 2014)
- It has around 181 texts and 3 638 coreferential chains

Смежные задачи

- Определение упоминаний (mention detection)
- Использование внешних знаний

Спасибо!

Вопросы?

Корпус RuCor: <http://rucoref.maimbava.net>

Репозиторий: <https://github.com/max-ionov/rucoref>

Тетрадки Jupyter:

<https://github.com/max-ionov/rucoref/tree/master/notebooks>