If we consider a discourse relation as a relationship between two phrases, we get a binary branching tree structure for the discourse. In many relationships, such as Explanation, one phrase depends on the other: e.g., the phrase being explained is the main one and the other is subsidiary. In fact we can get rid of the subsidiary phrases and still have a reasonably coherent discourse.

Lecture 11: Discourse
1. Coherence
2. Anaphora (pronouns etc)
3. Algorithms for anaphora resolution

most slides are from Ann Copestake
Coherence
Document structure and discourse structure

Most types of document are highly structured, implicitly or explicitly:

- Scientific papers: conventional structure (differences between disciplines).
- News stories: first sentence is a summary.
- Blogs, etc etc

Three Generative, Lexicalised Models for Statistical Parsing

**Abstract:** In this paper we first propose [...] We then [...] Results on Wall Street Journal text show [...] 

1. **Introduction:** [...] In the remainder of this paper we describe the 3 models in section 2, discuss practical issues in section 3, give results in section 4, and give conclusions in section 5.

2. **The Three Parsing Models**
3. **Practical Issues**
4. **Results**
5. **Conclusions**
Document structure and discourse structure

Most types of document are highly structured, implicitly or explicitly:

- Scientific papers: conventional structure (differences between disciplines).
- News stories: first sentence is a summary.
- Blogs, etc etc

US Election 2020: All eyes on Florida in Trump-Biden battle

Early results from the US presidential election, pitting Donald Trump against Joe Biden, show a tight race in the potentially pivotal race of Florida.

With 73% of votes counted in Florida, Mr Biden, the Democratic challenger, had a lead of 51% to 48% for Mr Trump, the Republican incumbent.

Polls have closed in three other key states, Georgia, Ohio and North Carolina, which are too close to call.

The vote caps a long and bitter campaign amid the coronavirus pandemic. More than 100 million people had already cast their ballots in early voting.
Max fell.       John pushed him.
1. Max fell *because* John pushed him.
1. Max fell *because* John pushed him.  ▷ EXPLANATION
2. Max fell *and then* John pushed him.  ▷ NARRATION
1. Max fell *because* John pushed him.  
2. Max fell *and then* John pushed him.

Rhetorical Structure Theory

Analysis of text with rhetorical relations generally gives a binary branching structure. Relations are defined between *elementary discourse unit (EDU)/discourse segments*, which can be a sentence, clause or phrase.

- **nucleus**: more central unit to the writer’s purpose that is interpretable independently
- **satellite**: less central unit that is only interpretable with respect to the nucleus.
- **equal weight**: e.g., NARRATION, LIST, CONTRAST
A large text is organized with a hierarchical structure

*With its distant orbit—50 percent farther from the sun than Earth—and slim atmospheric blanket, Mars experiences frigid weather conditions.* Surface temperatures typically average about -60 degrees Celsius (-76 degrees Fahrenheit) at the equator and can dip to -123 degrees C near the poles. Only the midday sun at tropical latitudes is warm enough to thaw ice on occasion, but any liquid water formed in this way would evaporate almost instantly because of the low atmospheric pressure.
Coherence in interpretation

Discourse coherence assumptions can affect interpretation

*Kim’s bike got a puncture. She phoned the AA.*

Assumption of coherence (and knowledge about the AA) leads to bike interpreted as motorbike rather than pedal cycle.

*John likes Bill. He gave him an expensive Christmas present.*

- If EXPLANATION - *he* is probably *Bill*.
- If JUSTIFICATION (supplying evidence for first sentence), *he* is *John*.
Goal: shortening long pieces of text

*If we consider a discourse relation as a relationship between two phrases, we get a binary branching tree structure for the discourse. In many relationships, such as EXPLANATION, one phrase depends on the other: e.g., the phrase being explained is the main one and the other is subsidiary. In fact we can get rid of the subsidiary phrases and still have a reasonably coherent discourse.*
Summarisation by satellite removal

**Goal:** shortening long pieces of text

If we consider a discourse relation as a relationship between two phrases, we get a binary branching tree structure for the discourse. In many relationships, such as **EXPLANATION**, one phrase depends on the other: e.g., the phrase being explained is the main one and the other is subsidiary. In fact we can get rid of the subsidiary phrases and still have a reasonably coherent discourse.
Coherence in generation

Language generation needs to maintain coherence

*In trading yesterday:* Dell was up 4.2%, Safeway was down 3.2%, HP was up 3.1%.

**Better:**

*Computer manufacturers gained in trading yesterday:* Dell was up 4.2% and HP was up 3.1%. *But retail stocks suffered:* Safeway was down 3.2%.

More about generation in the next lecture.
Factors influencing discourse interpretation

1. Cue phrases
   - *because*, *in order to*

2. Punctuation (also prosody) and text structure.
   - Max fell *(John pushed him)* and Kim laughed.
   - Max fell, John pushed him and Kim laughed.

3. Real world content:
   - Max fell. John pushed him as he lay on the ground.

4. Tense and aspect.
   - Max fell. John had pushed him.
   - Max was falling. John pushed him.

Hard problem, but ‘surfacy techniques’ (punctuation and cue phrases) work to some extent.
Discourse structure parsing

Two subtasks

• EDU segmentation
• RST parsing
Anaphora (pronouns etc)
Referring expressions

*Niall Ferguson* is prolific, well-paid and a snappy dresser. *Stephen Moss* hated *him* — at least until *he* spent an hour being charmed in *the historian’s* Oxford study.

**referent**: a real world entity that some piece of text (or speech) refers to, e.g. the actual Prof. Ferguson

**referring expressions/mentions**: bits of language used to perform reference by a speaker, e.g. *Niall Ferguson, he, him, the historian*
Referring expressions

*Niall Ferguson* is prolific, well-paid and a snappy dresser. *Stephen Moss* hated *him* — at least until *he* spent an hour being charmed in the historian’s *Oxford* study.

**referent**: a real world entity that some piece of text (or speech) refers to, e.g. the actual Prof. Ferguson

**referring expressions/mentions**: bits of language used to perform reference by a speaker, e.g. *Niall Ferguson, he, him, the historian*

**corefer**: two or more referring expressions refer to the same *discourse entity*
Referring expressions

*Niall Ferguson* is prolific, well-paid and a snappy dresser. *Stephen Moss* hated *him* — at least until *he* spent an hour being charmed in the historian’s Oxford study.

**referent**: a real world entity that some piece of text (or speech) refers to, e.g. the actual Prof. Ferguson

**referring expressions/mentions**: bits of language used to perform reference by a speaker, e.g. *Niall Ferguson, he, him, the historian*

**corefer**: two or more referring expressions refer to the same *discourse entity*

**antecedent**: the text initially evoking a referent, e.g. *Niall Ferguson*

**anaphora**: the phenomenon of referring to an antecedent.
Referring expressions

*Niall Ferguson* is prolific, well-paid and a snappy dresser. *Stephen Moss* hated *him* — at least until *he* spent an hour being charmed in *the historian’s* Oxford study.

**referent**: a real world entity that some piece of text (or speech) refers to, e.g. the actual Prof. Ferguson

**referring expressions/mentions**: bits of language used to perform reference by a speaker, e.g. *Niall Ferguson, he, him, the historian*

**corefer**: two or more referring expressions refer to the same *discourse entity*

**antecedent**: the text initially evoking a referent, e.g. *Niall Ferguson*

**anaphora**: the phenomenon of referring to an antecedent.

**coreference resolution**: the task of determining whether two mentions corefer
Referring expressions

*Niall Ferguson* is prolific, well-paid and a snappy dresser. *Stephen Moss* hated him — at least until *he* spent an hour being charmed in the *historian’s* Oxford study.

**referent**: a real world entity that some piece of text (or speech) refers to, e.g. the actual Prof. Ferguson

**referring expressions/mentions**: bits of language used to perform reference by a speaker, e.g. *Niall Ferguson, he, him, the historian*

**corefer**: two or more referring expressions refer to the same *discourse entity*

**antecedent**: the text initially evoking a referent, e.g. *Niall Ferguson*

**anaphora**: the phenomenon of referring to an antecedent.

**coreference resolution**: the task of determining whether two mentions corefer

**entity linking/entity resolution**: the task of mapping a mention in text to the representation of some real-world entity in an *ontology*, e.g. Wikipedia pages (wikification)
Event coreference

Event variable

(1)  a. Rover barked loudly
    b. $\exists e [\text{bark}'(e, r) \land \text{loud}'(e)]$

AMD agreed to buy Markham, Ontario-based ATI for around $5.4 billion in cash and stock, the companies announced Monday. The acquisition would turn AMD into one of the world’s largest providers of graphics chips.

event coreference: the task of deciding whether two event mentions refer to the same event
Hard constraints

Pronoun agreement (PERSON/NUMBER/GENDER)

(2) a. *A little girl is at the door — see what she wants, please?
   b. My dog has hurt *his* foot — *he* is in a lot of pain.
   c. *My dog has hurt his foot — it is in a lot of pain.

Complications

(3) a. *The team played really well, but now they are all very tired.
   b. *Kim and Sandy are asleep: they are very tired.
   c. *Kim is snoring and Sandy can’t keep her eyes open: they are both exhausted.

Reflexives

(4) a. John\textsubscript{i} cut himself\textsubscript{i} shaving. (*himself = John, subscript as ID*)
   b. #John\textsubscript{i} cut him\textsubscript{j} shaving. (*i \neq j — a very odd sentence*)

Reflexive pronouns must be coreferential with a preceeding argument of the same verb, non-reflexive pronouns cannot be.
Non-referring expressions

*pleonastic/expletive pronouns*: semantically empty, don’t refer
(5) a. It is snowing  
   b. It is not easy to think of good examples.  
   c. It is obvious that Kim snores.  
   d. It bothers Sandy that Kim snores.

*appositional NPs*: supplementary parenthetical descriptions

*predicative and Prenominal NPs*: properties of the head nouns
(7) the 38-year-old became *the company’s president*

*generics*: classes rather than individuals or sets of individuals
(8) I love *mangos*. *They* are very tasty.
Recency
Kim has a big car. Sandy has a smaller one. Lee likes to drive it.

Grammatical role: subjects $\succ$ objects $\succ$ everything else
Fred went to the Grafton Centre with Bill. He bought a hat.

Repeated mention: Entities that have been mentioned more frequently are preferred.

Parallelism: Entities which share the same role as the pronoun in the same sort of sentence are preferred
Bill went with Fred to the Grafton Centre. Kim went with him to Lion Yard.
World knowledge and common sense reasoning

Winograd schema challenge

(9) a. The city council denied the demonstrators a permit because
   b. they feared violence. ▷ they = the city council
   c. they advocated violence. ▷ they = the demonstrators

Sometimes inference will override soft preferences:

Andrew Strauss again blamed the batting after England lost to Australia last night. They now lead the series three-nil.

they = Australia

But a discourse can be odd if strong salience effects are violated:

The England football team won last night. Scotland lost.

They have qualified for the World Cup with a 100% record.
Algorithms for Anaphora Resolution
Anaphora resolution as supervised classification

- **Classification**: training data labelled with class and features, derive class for test data based on features.
- For potential pronoun/antecedent pairings, class is TRUE/FALSE.
- Assume candidate antecedents are all NPs in current sentence and preceeding 5 sentences (excluding pleonastic pronouns)

Example

*Niall Ferguson is prolific, well-paid and a snappy dresser. Stephen Moss hated him — at least until he spent an hour being charmed in the historian’s Oxford study.*

**Issues**: detecting pleonastic pronouns and predicative NPs, deciding on treatment of possessives (*the historian* and *the historian’s Oxford study*), named entities (e.g., *Stephen Moss*, not *Stephen* and *Moss*), allowing for cataphora, . . .
Features

*Cataphoric*: true if pronoun before antecedent.

*Number agreement*: true if pronoun compatible with antecedent.

*Gender agreement*: true if gender agreement.

*Same verb*: true if the pronoun and the candidate antecedent are arguments of the same verb.

*Sentence distance*: \{0, 1, 2, \ldots\} (discrete)

*Grammatical role* (the role of the potential antecedent): \{subject, object, other\}

*Parallel*: true if the potential antecedent and the pronoun share the same grammatical role.

*Linguistic form*: \{proper, definite, indefinite, pronoun\}
Problems with simple classification model

- Cannot implement ‘repeated mention’ effect.
- Cannot use information from previous links (not a structured prediction solution)
- Not really pairwise: really need discourse model with real world entities corresponding to clusters of referring expressions.
Zero anaphors

zero anaphor/zero pronoun: in some languages (including Chinese, Japanese, and Italian) it is possible to have an anaphor that has no lexical realization at all.

EN John\textsubscript{i} went to visit some friends. On the way he\textsubscript{i} bought some wine.

IT Giovanni\textsubscript{i} andò a far visita a degli amici. Per via $\emptyset\textsubscript{i}$ comprò del vino.

JA John\textsubscript{i}-wa yujin-o houmon-sita. Tochu-de $\emptyset\textsubscript{i}$ wain-o ka-tta.
Readings

• Ann’s lecture notes.
  https://www.cl.cam.ac.uk/teaching/1920/NLP/materials.html

• D Jurafsky and J Martin. *Speech and Language Processing*. Chapter 22. Coreference resolution.