Foundations of Natural Language Processing
Lecture 17
Discourse Coherence

Alex Lascarides

17 March 2020

School of informatics

Making sense of actions

Changing our minds

Observing Action

- We assume action choice isn't arbitrary;
  choice is informed by the context
- So we infer more than we see.
- And may change these inferences as we see more.

Coherence in Discourse: Making sense of verbal actions

It's a beautiful night.
We're looking for something dumb to do.
Hey baby, I think I wanna marry you.

Questions

Coherence and Content

Representation: How should discourse coherence be represented formally and computationally?

Construction: What inference processes, and what knowledge sources, are used when identifying coherence relations?

Outline

- Motivation for discourse coherence
- Representing discourse coherence
- Inferring discourse coherence

Pronouns

From Hobbs (1985)

John can open Bill's safe.
He knows the combination

Pronouns

From Hobbs (1985)

John can open Bill's safe.
John ~~He~~ knows the combination.

- If "He" is John: Explanation ("because").

Pronouns

From Hobbs (1985)

John can open Bill's safe.
Bill ~~He~~ knows the combination.

- If "He" is John: Explanation ("because").
  If "He" is Bill: at best we infer Continuation ("and") with a very vague topic.

Pronouns

From Hobbs (1985)

John can open Bill's safe.
He should change the combination.

Pronouns

From Hobbs (1985)

John can open Bill's safe.
Bill ~~He~~ should change the combination.

- If "He" is Bill: Result ("so")

Pronouns

From Hobbs (1985)

> John can open Bill's safe.
> John ~~He~~ should change the combination.

- If "He" is Bill: Result ("so")
  If "He" is John: a 'weaker' Result?

- Subjects are more likely antecedents, but not here...

Pronouns and Coherence

- Pronouns interpreted in a way that maximises coherence, even if this conflicts with predictions from other knowledge sources!

Coherence and Time

Max fell. John helped him up.
Max fell. John pushed him.

Coherence and Time

John hit Max on the back of his neck.
Max fell. John pushed him.
Max rolled over the edge of the cliff.

Word Meaning

A: Did you buy the apartment?
B: Yes, but we rented it./ No, but we rented it.

Bridging

John took an engine from Avon to Dansville.
He picked up a boxcar./He also took a boxcar.

Discourse Coherence and Implicit Agreement

From Sacks *et al.* (1974):

(1) a. M (to K and S): Karen 'n' I're having a fight,
b. M (to K and S): after she went out with Keith and not me.
c. K (to M and S): Wul Mark, you never asked me out.

Discourse Coherence and Dishonesty

Example from Solan and Tiersma (2005)

(2) a. P: Do you have any bank accounts in Swiss banks, Mr. Bronston?
b. B: No, sir.
c. P: Have you ever?
d. B: The company had an account there for about six months, in Zurich.

- (2)d interpreted as an indirect answer, implying *no*. . .

Discourse Coherence and Dishonesty

Example from Solan and Tiersma (2005)

(2)
a. P: Do you have any bank accounts in Swiss banks, Mr. Bronston?
b. B: No, sir.
c. P: Have you ever?
d. B: The company had an account there for about six months, in Zurich.

- (2)d interpreted as an indirect answer, implying *no*. . .
- . . . *even if* you know it conflicts with Bronston's beliefs.

Discourse Coherence and Dishonesty

Example from Solan and Tiersma (2005)

(2)
a. P: Do you have any bank accounts in Swiss banks, Mr. Bronston?
b. B: No, sir.
c. P: Have you ever?
d. B: The company had an account there for about six months, in Zurich.

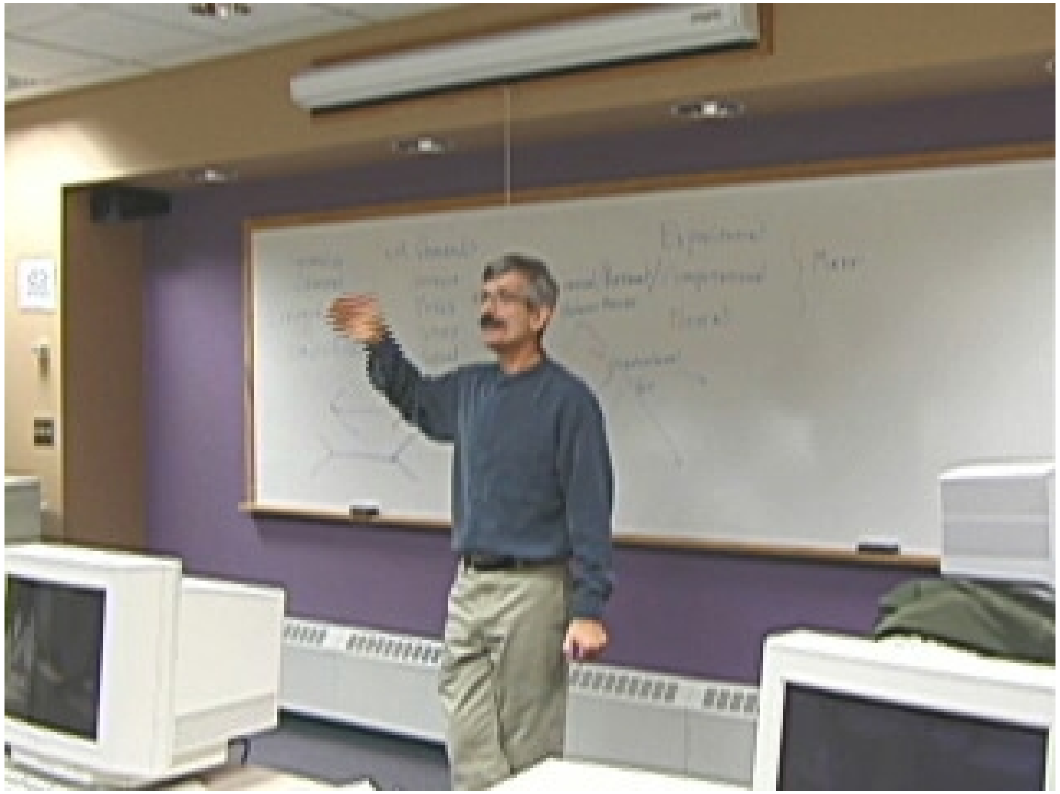
- (2)d interpreted as an indirect answer, implying *no*. . .
- . . . *even if* you know it conflicts with Bronston's beliefs.
- Literally true, but negative answer false.
- Supreme court overruled conviction for perjury.
- Different ruling probable if Bronston had said "only".

Gesture

Now one thing you could do is totally audiotape hours and hours. . .

. . . so that you get a large amount of data that you can think of as laid out on a time line.

Gesture



And exhaustively go through and make sure that you really pick up all the speech errors

. . . by individually analysing each acoustic unit along the timeline of your data.

Gesture

Allow two different coders to go through it. . .

. . . and moreover get them to work independently and reconcile their activities.

Gesture and Coherence

Lascarides and Stone (2009)

Meaning of Multimodal Communicative Actions

Coherence relations connect speech and gesture and sequences of gestures.

- speech so that gesture
  speech by gesture
  speech and moreover gesture

SDRT: The logical form (LF) of monologue

LF consists of:

1. Set A of labels $\pi_1, \pi_2, \ldots$
   (each label stands for a segment of discourse)

2. A mapping $\mathcal{F}$ from each label to a formula representing its content.

3. Vocabulary includes coherence relations; e.g., *Elaboration*(π_1, π_2).

LFs and Coherence

Coherent discourse is a single segment of rhetorically connected subsegments. More formally:

- The partial order over A induced by $\mathcal{F}$ has a unique root.

An Example

π_1: John can open Bill's safe.
π_2: He knows the combination.

$$\begin{array}{ll} \pi_0: & \textit{Explanation}(\pi_1, \pi_2) \\ \pi_1: & \iota x(\textit{safe}(x) \ \& \ \textit{possess}(x, \textit{bill}) \ \& \ \textit{can}(\textit{open}(e_1, \textit{john}, x)) \\ \pi_2: & \iota y(\textit{combination}(y) \ \& \ \textit{of}(y, x) \ \& \ \textit{knows}(\textit{john}, y)) \end{array}$$

- Bits in red are specific values that go beyond content that's revealed by linguistic form.

- They are inferred via commonsense reasoning that's used to construct a maximally coherent interpretation.

SDRT: Logical form of dialogue Lascarides and Asher (2009)

- LF tracks all current public commitments for each agent, including commitments to coherence relations.

(1) a. M (to K and S): Karen 'n' I're having a fight,
 b. M (to K and S): after she went out with Keith and not me.
 c. K (to M and S): Wul Mark, you never asked me out.

Turn	M	K
1	π_{1M} : *Explanation*(a, b)	$\emptyset$
2	π_{1M} : *Explanation*(a, b)	π_{2K} : *Explanation*$(a, b) \wedge$ *Explanation*(b, c)

Dishonesty

Asher and Lascarides (2011)

(2)
a. P: Do you have any bank accounts in Swiss banks?
b. B: No, sir.
c. P: Have you ever?
d. B: The company had an account there for 6 months.

Turn	Prosecutor	Bronston
1	$a : \mathcal{F}(a)$	$\emptyset$
2	$a : \mathcal{F}(a)$	π_{2B} : *Answer*(a, b)
3	π_{3P} : *Continuation*(a, c)	π_{2B} : *Answer*(a, b)
4	π_{3P} : *Continuation*(a, c)	π_{4B} : *Answer*$(a, b) \wedge$ *Continuation*$(a, c) \wedge$ *Indirect-Answer*(c, d)

1. Plausible Deniability: Must test rigorously whether it's safe to treat the implied answer as a matter of public record.

Dishonesty

Asher and Lascarides (2011)

(2)
a. P: Do you have any bank accounts in Swiss banks?
b. B: No, sir.
c. P: Have you ever?
d. B: The company had an account there for 6 months.

Turn	Prosecutor	Bronston
1	$a : \mathcal{F}(a)$	$\emptyset$
2	$a : \mathcal{F}(a)$	π_{2B} : *Answer*(a, b)
3	π_{3P} : *Continuation*(a, c)	π_{2B} : *Answer*(a, b)
4	π_{3P} : *Continuation*(a, c)	π_{4B} : *Answer*$(a, b) \wedge$ *Continuation*(b, d)

1. Plausible Deniability: Must test rigorously whether it's safe to treat the implied answer as a matter of public record.

2. Neologism proof equilibria: distinguishes (2)d vs. "only".

Symbolic approaches to constructing LF

- Draw on rich information sources:
  - linguistic content, world knowledge, mental states...
- Deploy reasoning that supports inference with partial information. Unlike classical logic, this requires consistency tests.
- Typically, construct LF and evaluate it in the same logic, making constructing LF undecidable.

Further Problem

- Like any knowledge rich approach involving hand-crafted rules, this is only feasible for very small domains.

- Ideally, we would like to learn a discourse parser automatically from corpus data.

- But there's a lack of corpora annotated with discourse structure.
  - RSTbank, Graphbank, Annodis, STAC are relatively small.
  - Discourse Penn Treebank is relatively large but not annotated with complete discourse structure.
  - Groningen Parellel Meaning Bank: full discourse structure (SDRSs) and getting bigger all the time.

Supervised Learning for SDRT

Training on 100 dialogues — Baldridge and Lascarides (2005)
Parser based on Collins' parsing model:

- 72% f-score on segmentation (baseline: 53.3%)
- 48% f-score on segmentation and coherence relations (baseline: 7.4%)
- Doesn't attempt to estimate LFs of clauses.

Training on Groningen Meaning Bank — Liu and Lapata (2018)
Neural semantic parser, RNN computes structure first, fills in arguments later:

- 77% f-score on segmentation, coherence relations *and* LFs of clauses
- State of the Art!

Avoiding Annotation

- Coherence relations can be overtly signalled:
  - *because* signals EXPLANATION; *but* signals CONTRAST
- So produce a training set *automatically*:
  - Max fell because John pushed him
    $\Rightarrow$
    EXPLANATION(*Max fell*, *John pushed him*).

Results of Best Model

- Test examples originally had a cue phrase: 60.9%.
- Test examples originally had no cue phrase: 25.8%
- Train on 1K manually labelled examples: 40.3%.
- Combined training set of manual and automatically labelled examples doesn't improve accuracy.

So you're better off manually labelling a small set of examples!

Results of Best Model

- Test examples originally had a cue phrase: 60.9%.
- Test examples originally had no cue phrase: 25.8%
- Train on 1K manually labelled examples: 40.3%.
- Combined training set of manual and automatically labelled examples doesn't improve accuracy.

So you're better off manually labelling a small set of examples!

Why?

Contrast to *Elaboration*
Although the electronics industry has changed greatly, possibly the greatest change is that very little component level manufacture is done in this country.

Conclusion

- Interpretation governed by discourse coherence:
  - Constrains what can be said next
  - Augments meaning revealed by linguistic form.
- Computing logical form should be decidable; modularity is key to this.
- Data-driven approaches are a major challenge.
- Linking rich models of discourse semantics to models of human behaviour and decision making is also a major challenge, but essential for tackling dialogues where the agents' goals conflict.