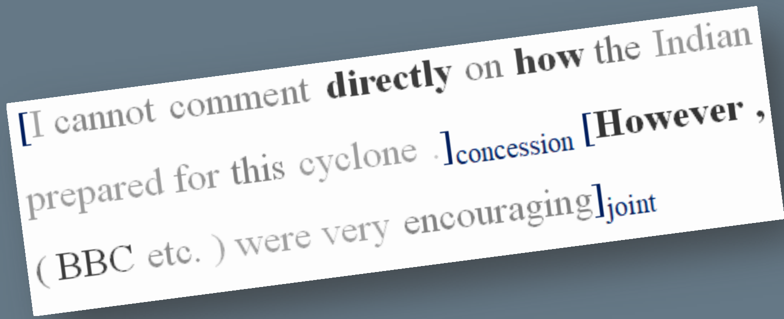
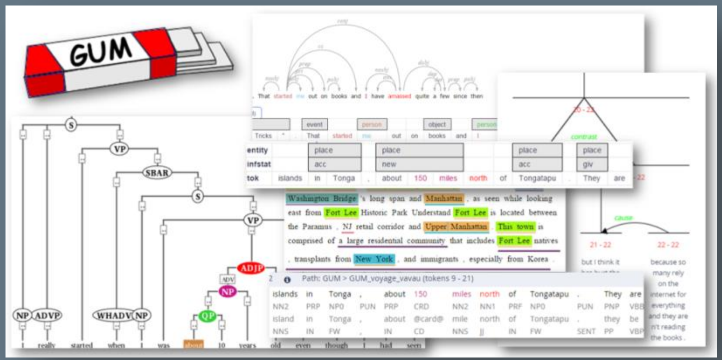


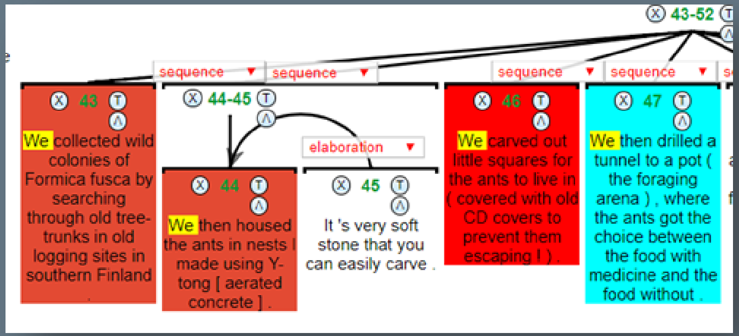


A Multilayer View of Discourse Relation Graphs

Amir Zeldes

Georgetown University

amir.zeldes@georgetown.edu



LTI Colloquium, CMU, 2017-11-10

Plan

I. Discourse parses in Rhetorical Structure Theory
   - What are discourse relations?
   - RST in a nutshell
   - Dependency representations for RST

II. Discourse encapsulation
   - Veins Theory and the Discourse Encapsulation Hypothesis
   - The Georgetown University Multilayer corpus (GUM)
   - A multifactorial model of discourse encapsulation

III. Relation signaling
   - RST-DT and the RST Signalling Corpus
   - Training LSTMs for signal detection

Discourse relations

- What relations exist between utterances as a text unfolds?

1. a. [[*John pushed Mary.*]$_{\text{cause}}$ *She fell.*]
   b. [[*Mary fell.* [*John pushed her.*]$_{\text{cause}}$]

(see Webber 1988, Asher & Lascarides 2003)

2. [[[*They left lights on*]$_{\text{cause}}$ *so Ellie got mad.*] [*She hates that*]$_{\text{background}}$]

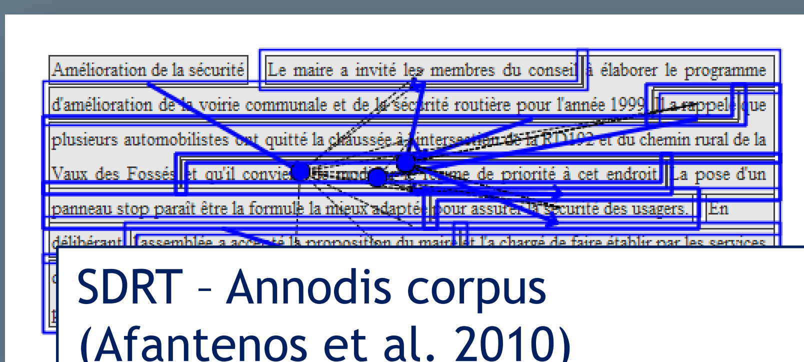
Discourse relations

- Some questions:
  - What relations exist? (Knott 1996, Knott & Sanders 1998)
    - Cross-linguistically? (van der Vliet & Radeker 2014)
    - In genres? (Taboada & Lavid 2003)
  - How are relations marked? (Taboada & Das 2013)
    - Explicit signals: *"on the other hand"* or *"although"*
    - Implicit signals: coreferent mentions, genre conventions, …
  - How do discourse relations constrain text organization? (Cristea et al. 1998, 1999; Tetreault & Allen 2003)

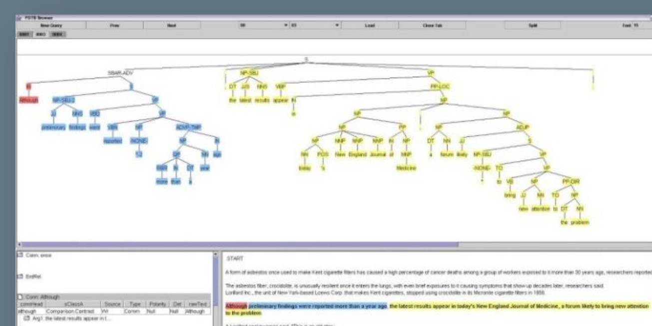
- To answer these questions we build **discourse annotated corpora**

Discourse annotation

- The task – given an arbitrary text:
  - Segment into ‘units’ (a.k.a. **E**lementary **D**iscourse **U**nits)
  - Establish the connections between these EDUs
  - Classify these connections
- Three main frameworks have implemented these tasks:
  - **Penn Discourse Treebank** (PDTB, Prasad et al. 2008) – partial parses
  - **Segmented Discourse Representation Theory** (SDRT, Asher & Lascarides 2003) – complete DAGs
  - **Rhetorical Structure Theory** (Mann & Thompson 1988) – complete trees



SDRT – Annodis corpus (Afantenos et al. 2010)



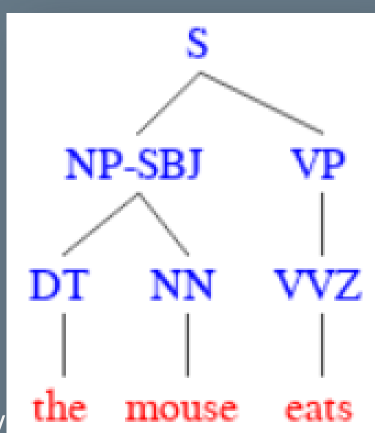
PDTB V2 (Prasad et al. 2008)

Rhetorical Structure Theory

- In RST, a text is a tree of clauses

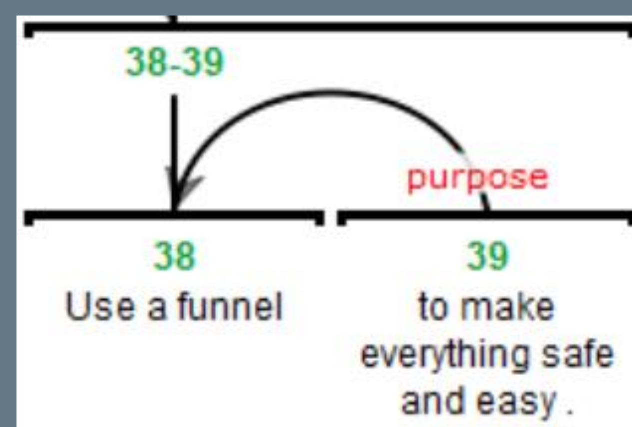
Syntax trees

- head > expansion
- Leaf = token
- Non-terminal = phrase
- Grammatical function

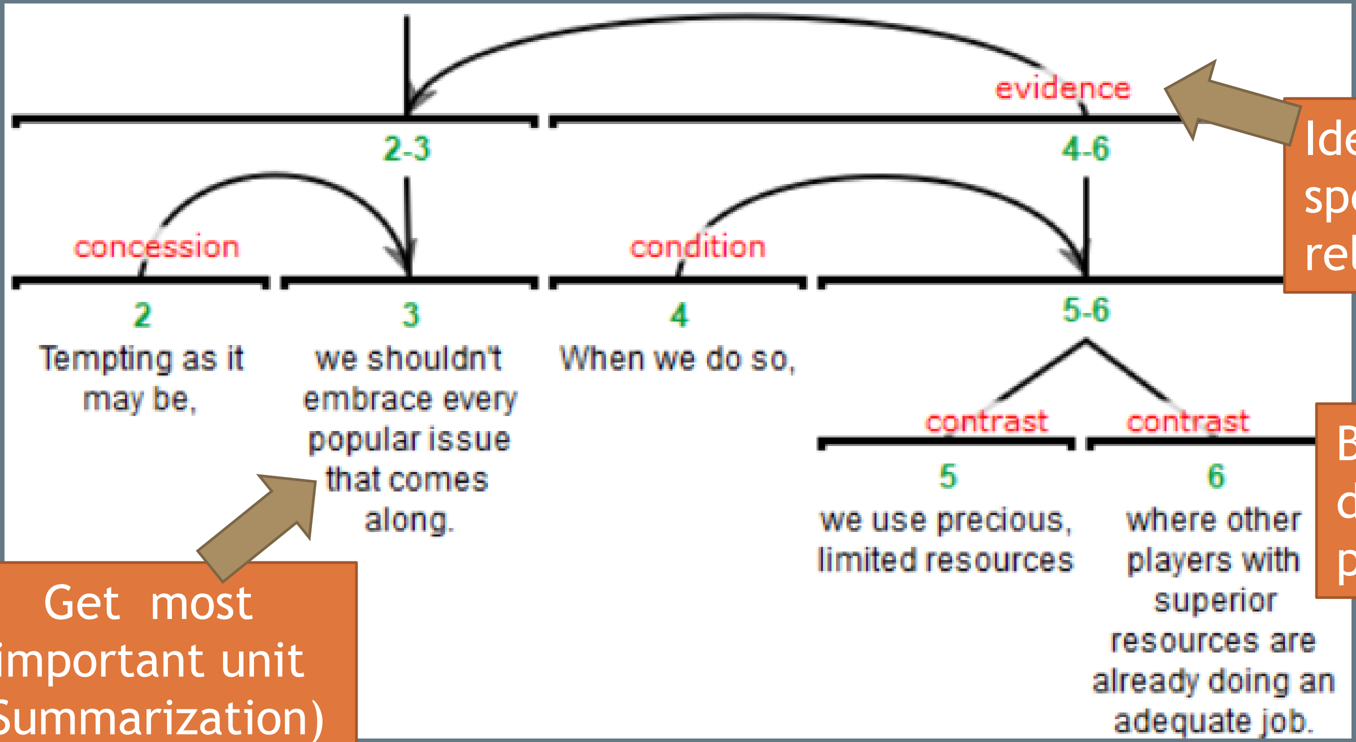


RST trees

- nucleus > satellite
- Leaf = EDU
- Non-terminal = span
- Discourse function



Why is this important?



Identify specific relations (IR)

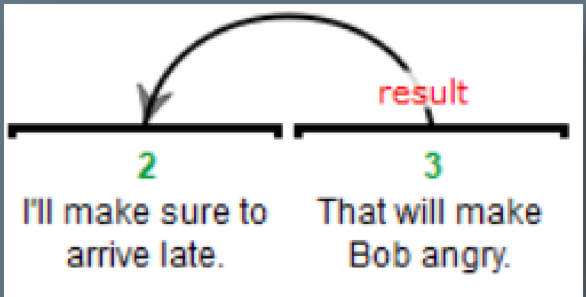
Build discourse plan (NLG)

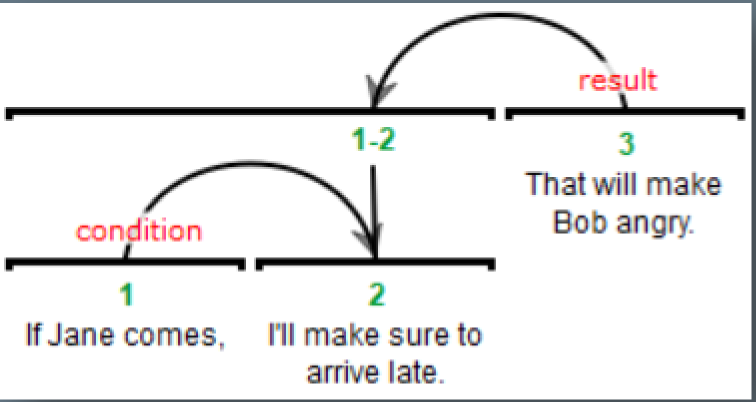
Get most important unit (Summarization)

(example from RST Website: http://www.sfu.ca/rst/)

Simplifying trees

- We will care how far things are in the graph
- Using non-terminal spans is problematic:





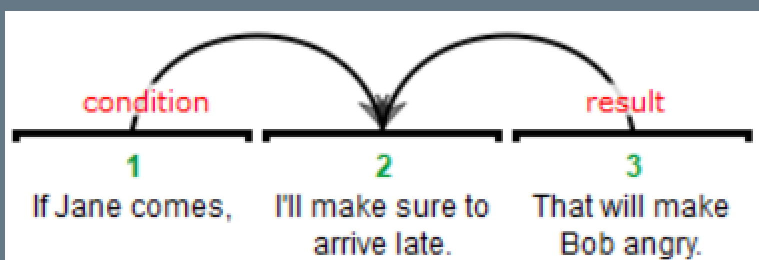
RD(2,3) = 1 edge

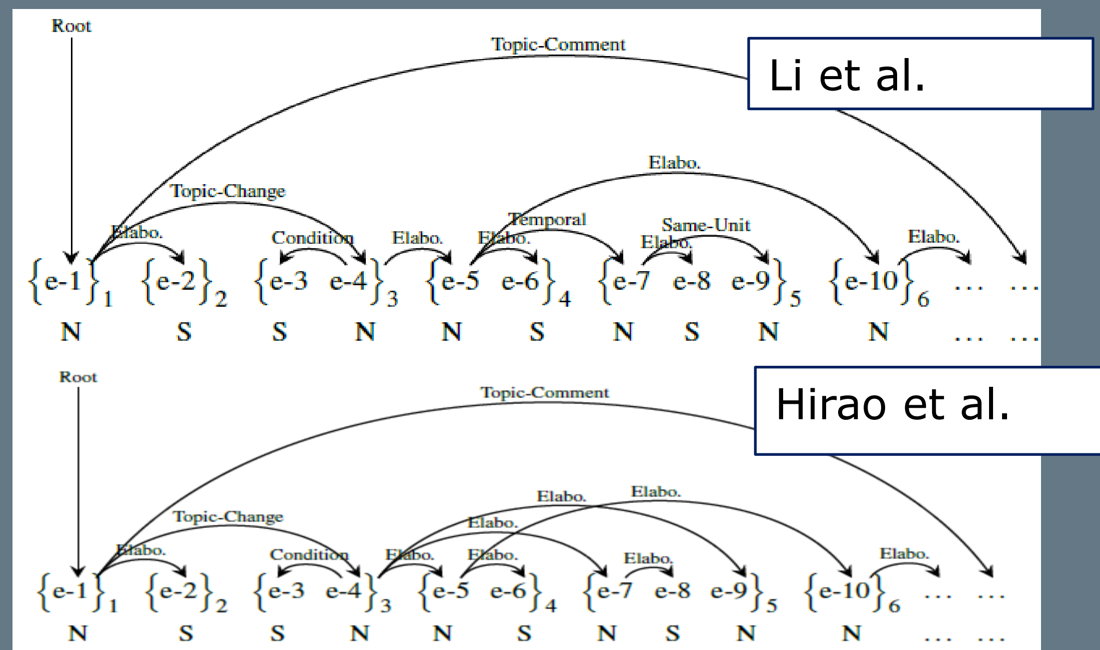
RD(2,3) = 2 edges

What's the "Rhetorical Distance"?

Dependency Representation

- Following Hayashi et al. (2016), use Li et al.'s (2014) dependency interpretation*





* conversion code available at: https://github.com/amir-zeldes/rst2dep

II. Discourse Encapsulation

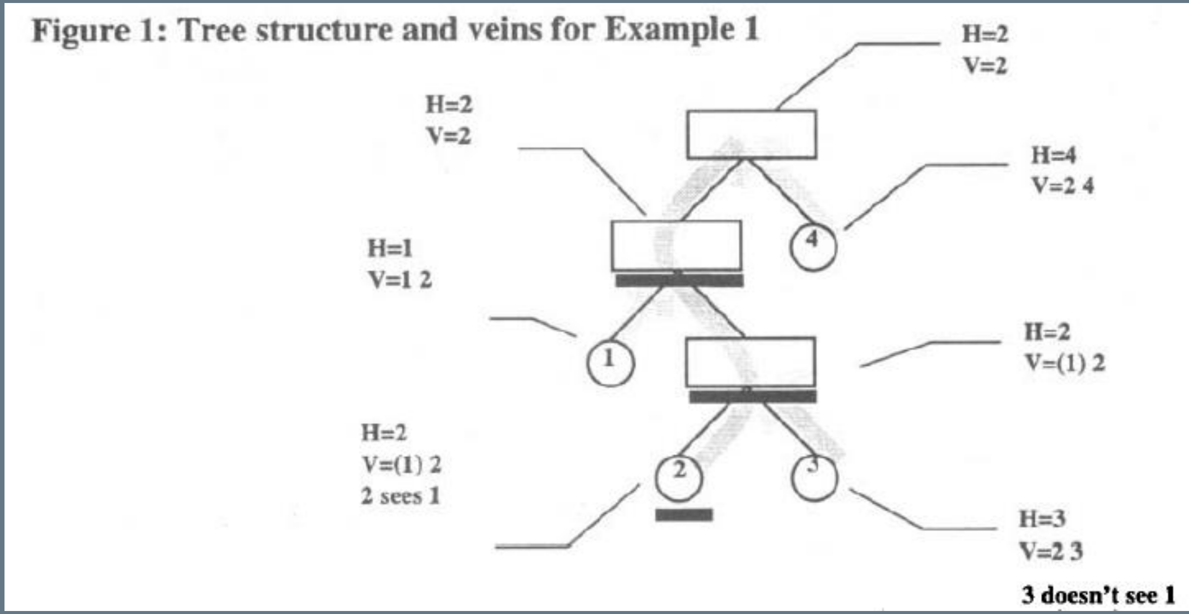
Discourse Encapsulation Hypothesis

- Do discourse parses constrain referentiality?
  - Discourse as stack (Polanyi 1988, Roberts 2012)
  - Right Frontier Constraint (Asher & Lascarides 2003)
  - Veins Theory (Cristea et al. 1998)
  - Different parametrizations (Chiarcos & Krasavina 2008)

Applications:
Coreference resolution
Referring expression generation
Dialog planning

Veins Theory (Cristea et al. 1998)

- VT proposes **Domains of Referential Accessibility**
  - Nuclei 'see' the satellites along their "vein"
  - Satellites can't access satellites of other nuclei
  - Path length irrelevant
  - Test on 5 texts: (fra, rom, eng) ~100%



Or not?

- Tetreault & Allen (2003:7):
  - *Our results indicate that **incorporating discourse structure** does not improve performance, and in most cases can actually **hurt performance**.*

- Based on much larger RST Discourse Treebank (RST-DT, ~180K tokens, Carlson et al. 2003)
- Suggests VT does not work 'in the wild'

Research questions

- Can we treat DRAs as quantitative tendencies?
  - Not binary restriction: more <--> less access
  - Multifactorial
    - Not just based on path
    - Also consider surface distance, graph distance, and more
  - Applicable to different types of referentiality?
    - Pronominal anaphora (*The president … he*)
    - Lexical coreference (*Joe Biden … Joe*, cf. TextTiling, Hearst 1997)
    - Bridging (*Mexico … the economy*; previously unstudied, cf. Asher & Lascarides 1998, Poesio & Vieira 1998, 2000)

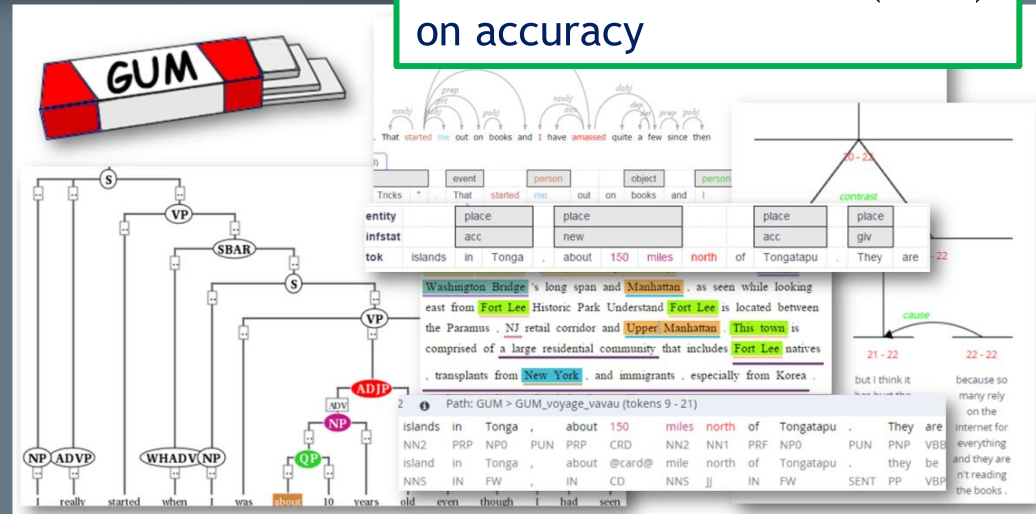
Data

- What could influence mention likelihood?
  (Recasens et al. 2013, Zeldes 2017a)
- We need:
  - RST parses
  - Coreference annotation (anaphora, lexical, bridging)
- Possible predictors:
  - Utterance length
  - Surface and ‘Rhetorical’ Distance metrics (SD, RD)
  - Syntactic structure (parses)
  - POS tags
  - Sentence types
  - …

Georgetown University Multilayer corpus

See Zeldes & Simonson (2016) on accuracy

- POS tagging (PTB, CLAWS, TT)
- Sentence type (SPAAC++)
- Document structure (TEI)
- Syntax trees (PTB + Stanford)
- Information status (SFB632)
- (Non-) named entity types
- **Coreference + bridging**
- **Rhetorical Structure Theory**
- Speaker information, ISO time…

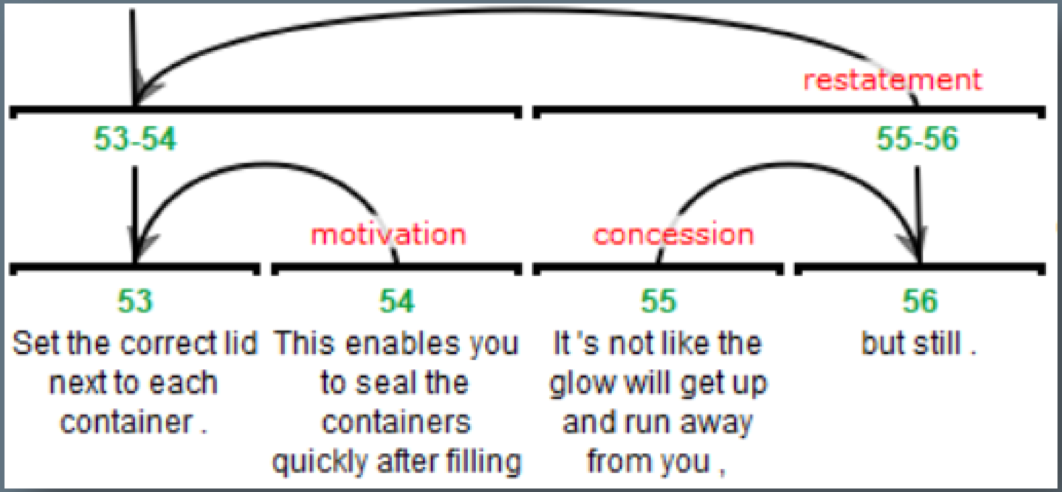


http://corpling.uis.georgetown.edu/gum/

text type	source	texts	tokens
Interviews (conversational)	Wikinews	19	18037
News (narrative)	Wikinews	21	14093
Travel guides (informative)	Wikivoyage	17	14955
How-tos (instructional)	wikiHow	19	16920
Total		76	64005

Veins in dependency representation

- Ancestry: Is one EDU a direct ancestor of the other in the dependency tree?



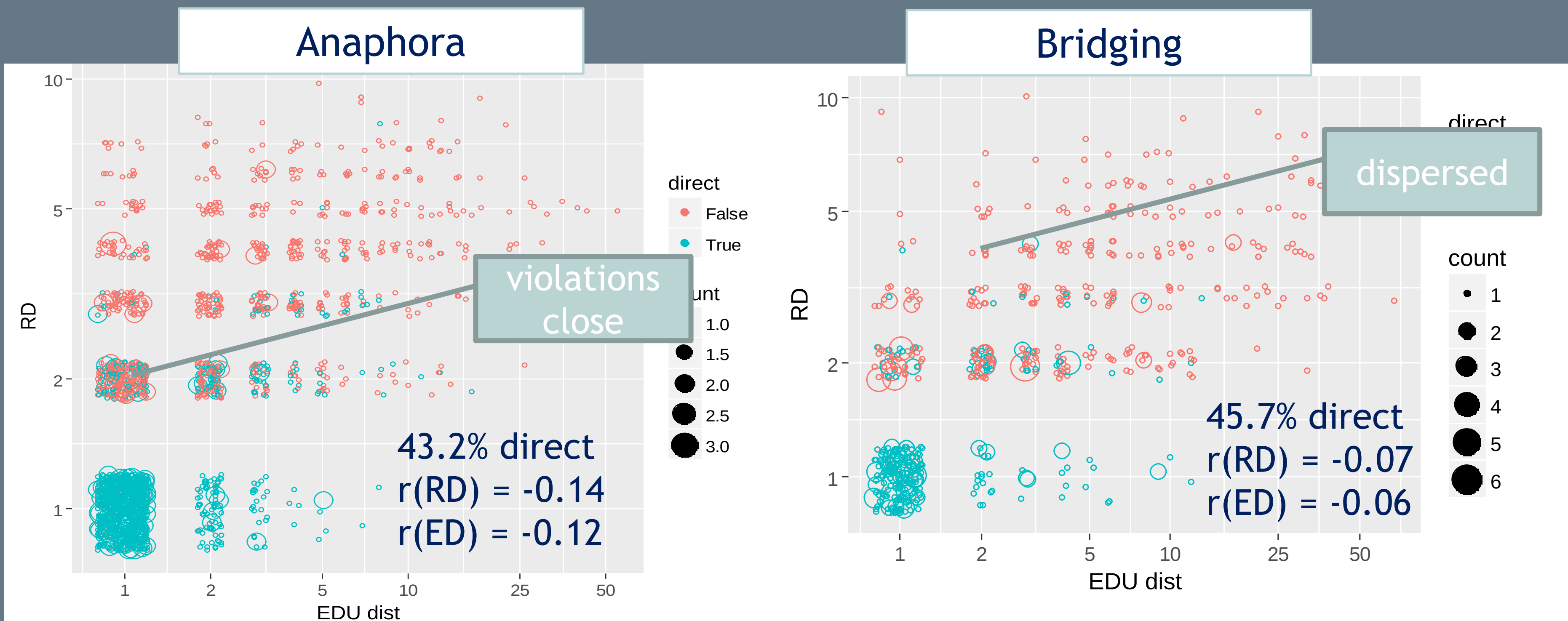
wikiHow:
"How to Make a Glowstick"

Target variables

- What are we trying to predict?
  - Binary domains:
    - Is there coreference between two EDUs?
    - Explore for anaphora, lexical, bridging
  - Coreference **density**:
    - How much coreferentiality exists between two EDUs? (# coreferent pairs)
  - Direct and indirect antecedents:
    - Check if the **immediate antecedent** of entity in EDU2 is in EDU1 (NB: makes surface distance very important!)
    - Alternatively, just check for coreference

Experiment setup

- ~170K possible EDU pairs grouped by document
- Looking at distance and direct parentage:



Why is prediction weak despite intuition?

Go multifactorial!

- Lots of confounds!!
  - **Length:** what if the main vein nucleus is really short? -> Unlikely to contain coreferent mentions
  - **Relations:** ***Purpose*** --> less coref; ***Cause*** --> more:
    - *needs to be exaggerated* [*in order to be funny —*]PURPOSE
    - *the banner read ' We Know ' .* [*That 's all it said .*]RESTATEMENT
  - **Sentence type:** imperatives, fragments have fewer entities than declaratives, questions
  - … + tense, genre, syntax, document position, …

Is RD significant? (any distance coref)

- Yes, and so is surface distance and directness!
- But not as important as length

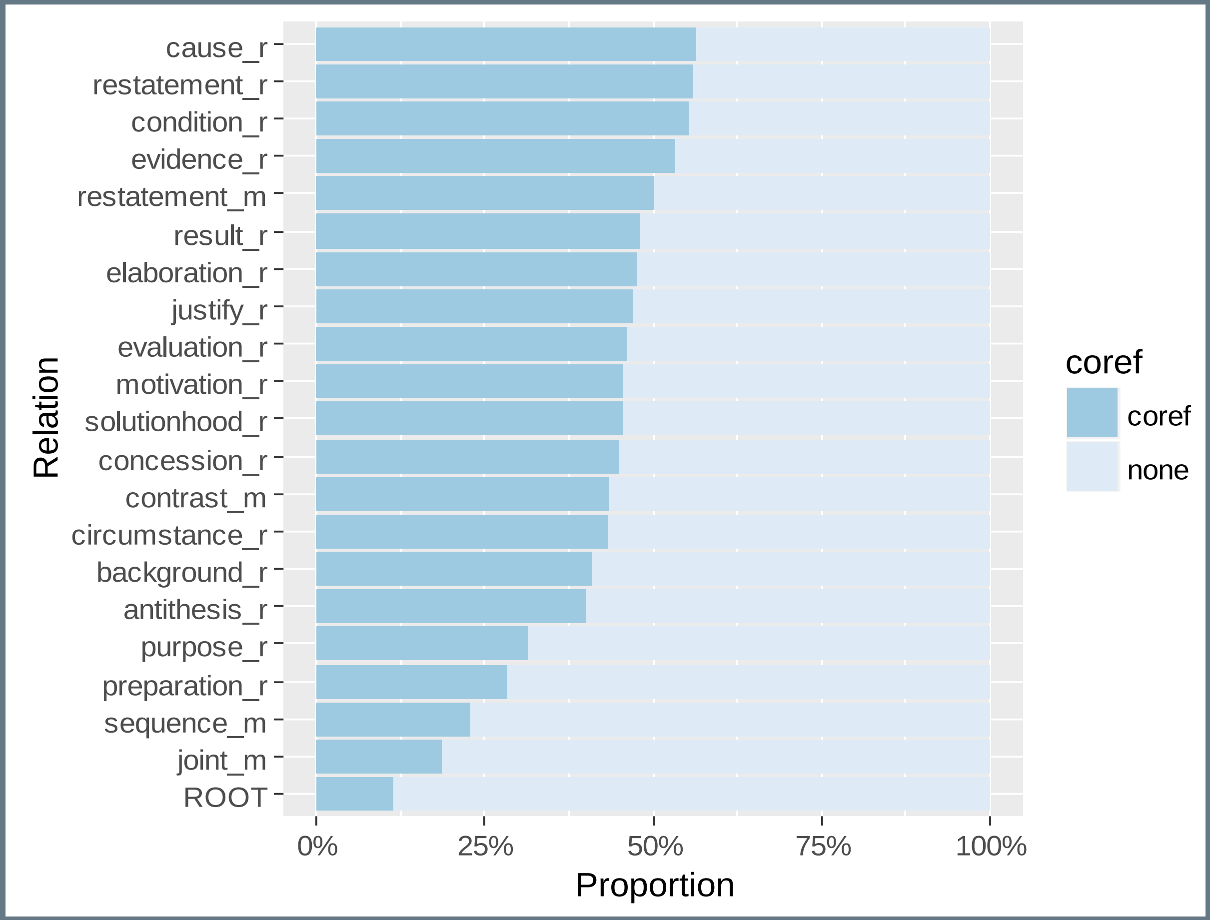
```
Random effects:
 Groups   Name        Variance Std.Dev.
 doc      (Intercept) 0.09789  0.3129
 Residual             0.82965  0.9109
Number of obs: 172150, groups:  doc, 76

Fixed effects:
              Estimate Std. Error t value
(Intercept)  0.2695836  0.0723038    3.73
scale(len1)  0.2043943  0.0023432   87.23 ***
scale(len2)  0.1833124  0.0023811   76.99 ***
rhet_dist   -0.0511588  0.0014351  -35.65 ***
edu_dist    -0.0015377  0.0001168  -13.17 ***
genrenews   -0.0348780  0.0997936   -0.35
genrevoyage -0.2161897  0.1047555   -2.06 **
genrewhow    0.0969725  0.1016942    0.95
directTrue   0.2280120  0.0091334   24.96 ***
```

Gaussian mixed effects model

Which relations favor coreference?

- Unsurprisingly:
  - ↑ Cause, Restatement
  - …
  - ↓ Joint, Sequence

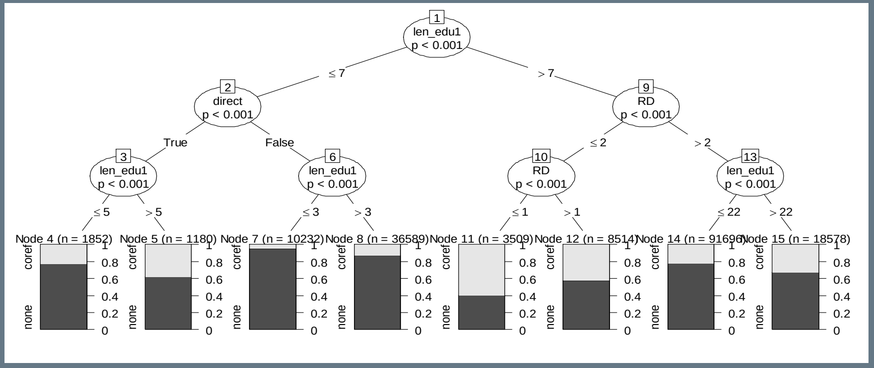


Add to linear model??

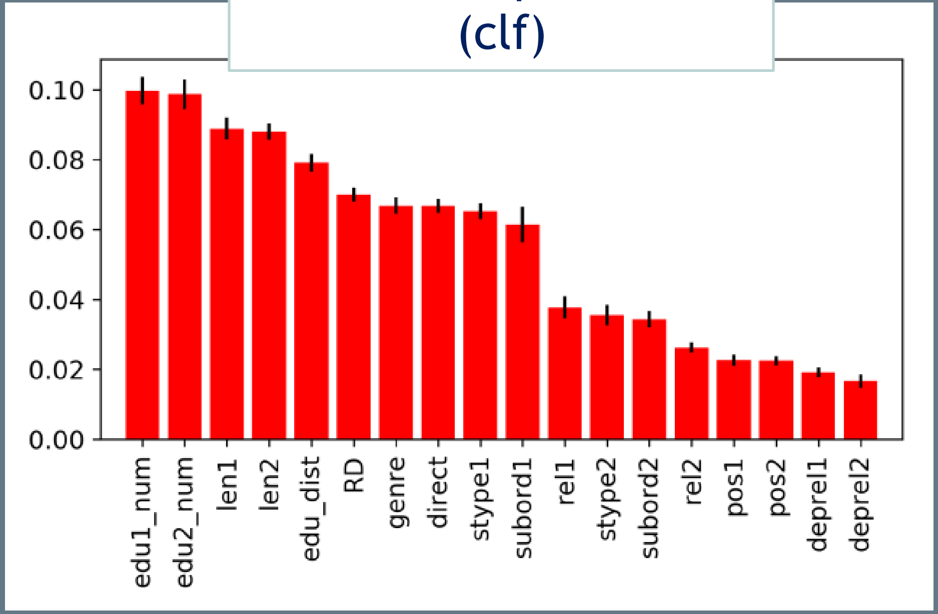
Ensemble approach (RST workshop@INLG 2017)

- Use Extra Trees ensemble (Geurts et al. 2006)
  - Classification (coref yes/no)
  - Regression (predict density)

features	RMSE (reg)	accuracy (clf)
majority	0.9652	77.90%
EDU	0.9501	78.36%
RD	0.9453	78.79%
all	0.7107	86.83%

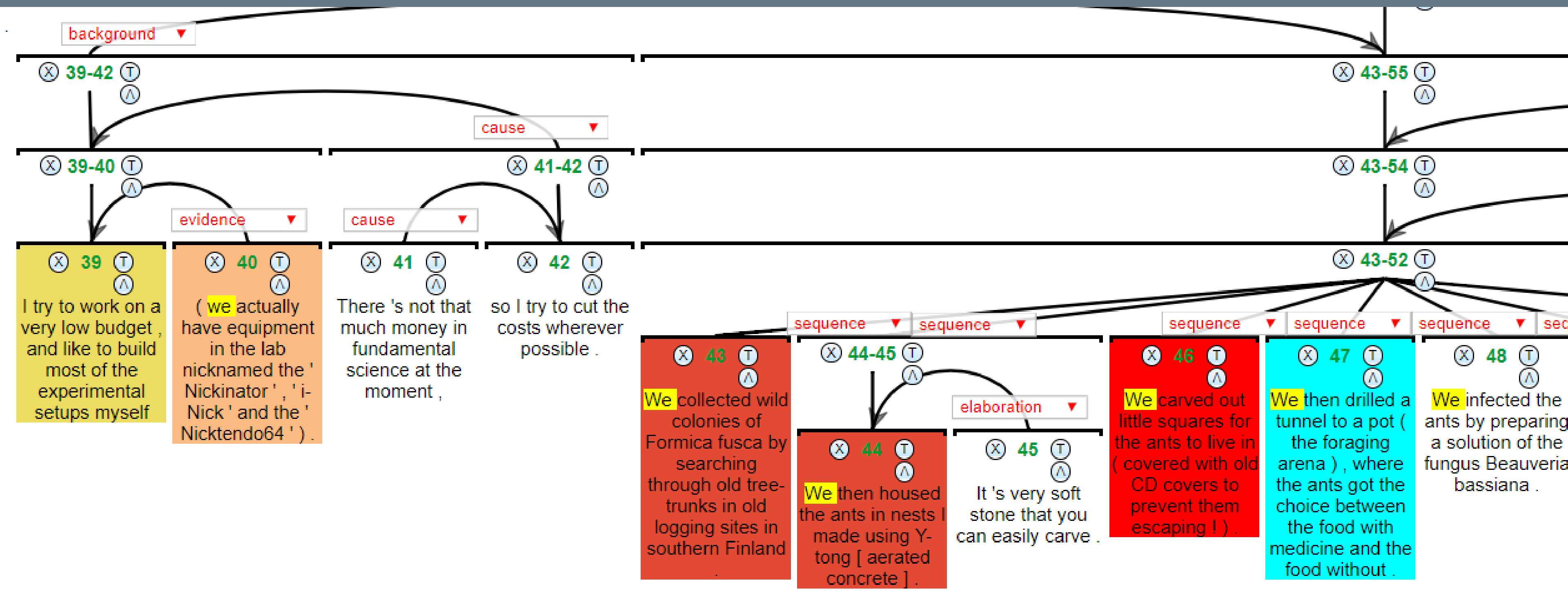


Feature importances (clf)

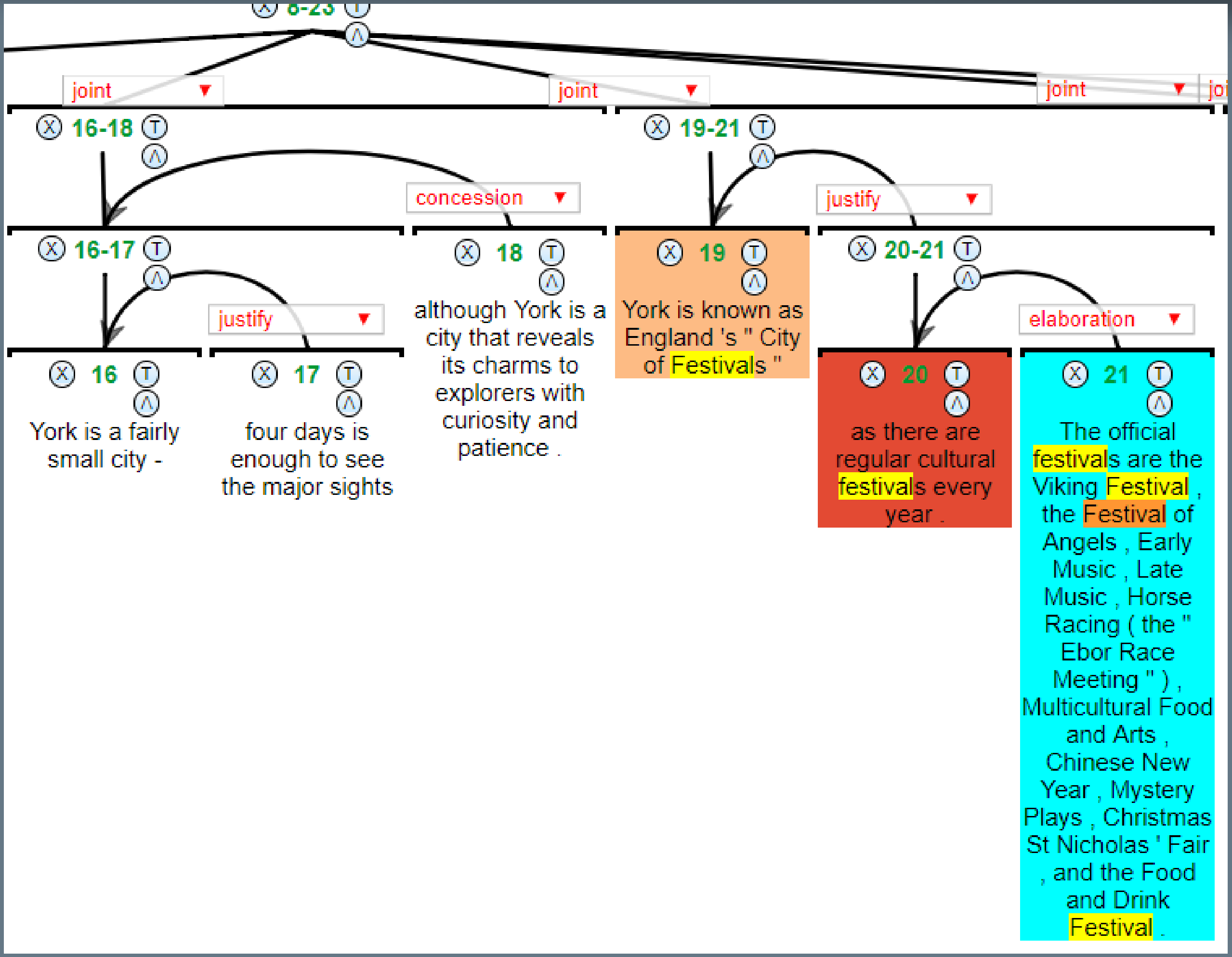


What do the predictions look like?

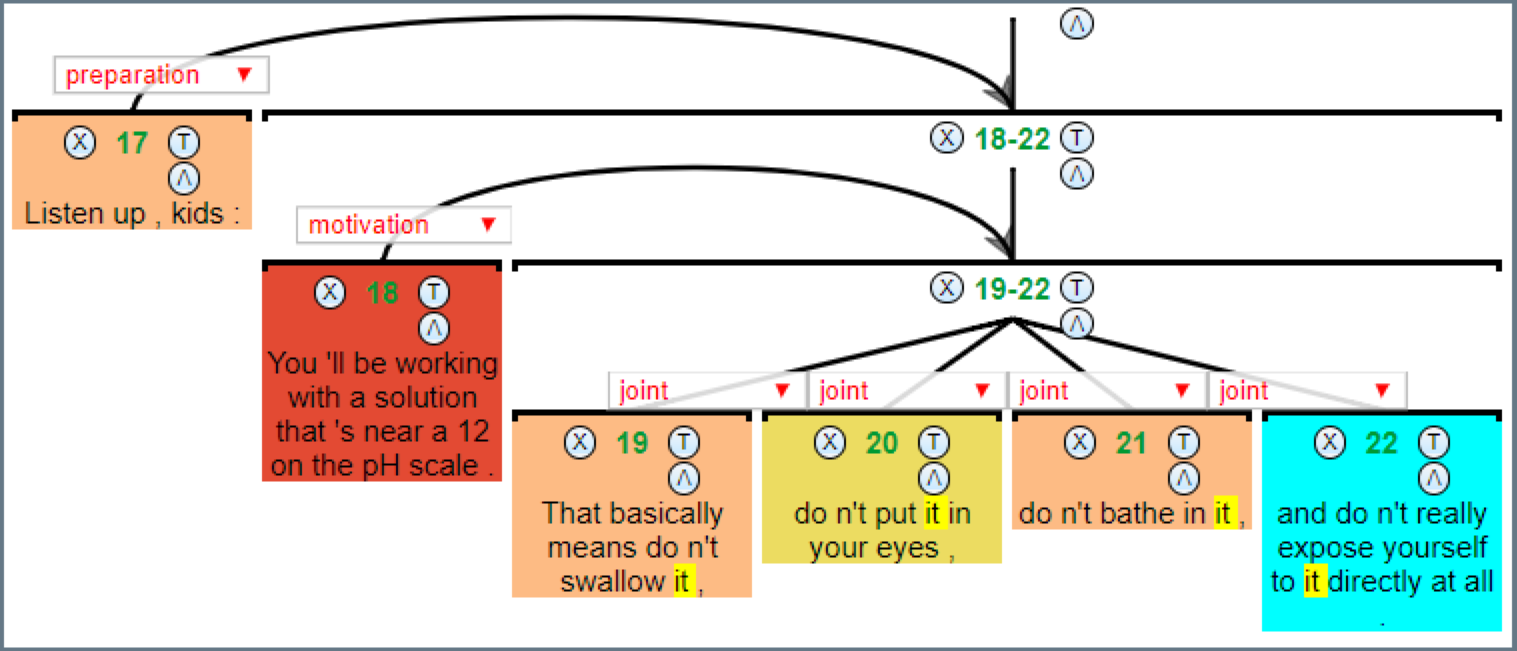
- We can visualize predictions as a heat map:



What do the predictions look like?



What do the predictions look like?



III. Relation signaling

Signaling

- Central question in discourse studies:
  - **Cues** help us to spot relations
  - Annotators use cue words as diagnostics:
    - *“could I connect these with ‘because’?”*
- Many approaches to relation taxonomies rely on discourse markers – connectives and other adverbials (Sweetser 1990, Sanders et al. 1992, Knott & Dale 1994, Taboada & Lavid 2003, Stede & Grishina 2016)

Research questions

- What kinds of signals are there?
- How can we identify them in data?
  - Are signal words always meaningful?
  - How ambiguous are they?
  - Can we distinguish meaningful and non-meaningful uses of cues?

Frequentist approaches

- Studies often cross-tabulate: ***words ~ relations***
- Problems:
  - Frequency thresholds
  - Ambiguity ("and" is not associated with any relation – not a Discourse Marker?)
  - Context sensitivity – some words are cues in specific environments

Relation type	Freq	marker	translation
Elaboration	150	*kotoryj*	"which, that"
Joint	119	*i, takzhe*	and, as well
Attribution	118	*zajavil, soobschil*	report, an-nounce etc.
Contrast	62	*Odnako, a, no*	However, but
Cause-Effect	47	*Poetomu, V+prichina*	so, accord-ingly, V+*cause*
Purpose	39	*Chtoby, dlya*	In order that, for
Interpretation-Evaluation	34	Nouns and verbs ex-pressing opinion	
Background	31	No domi-nant mark-er	
Condition	27	*esli*	if

Table 1. Relations with their most frequent markers

Toldova et al. 2017

Example – GUM

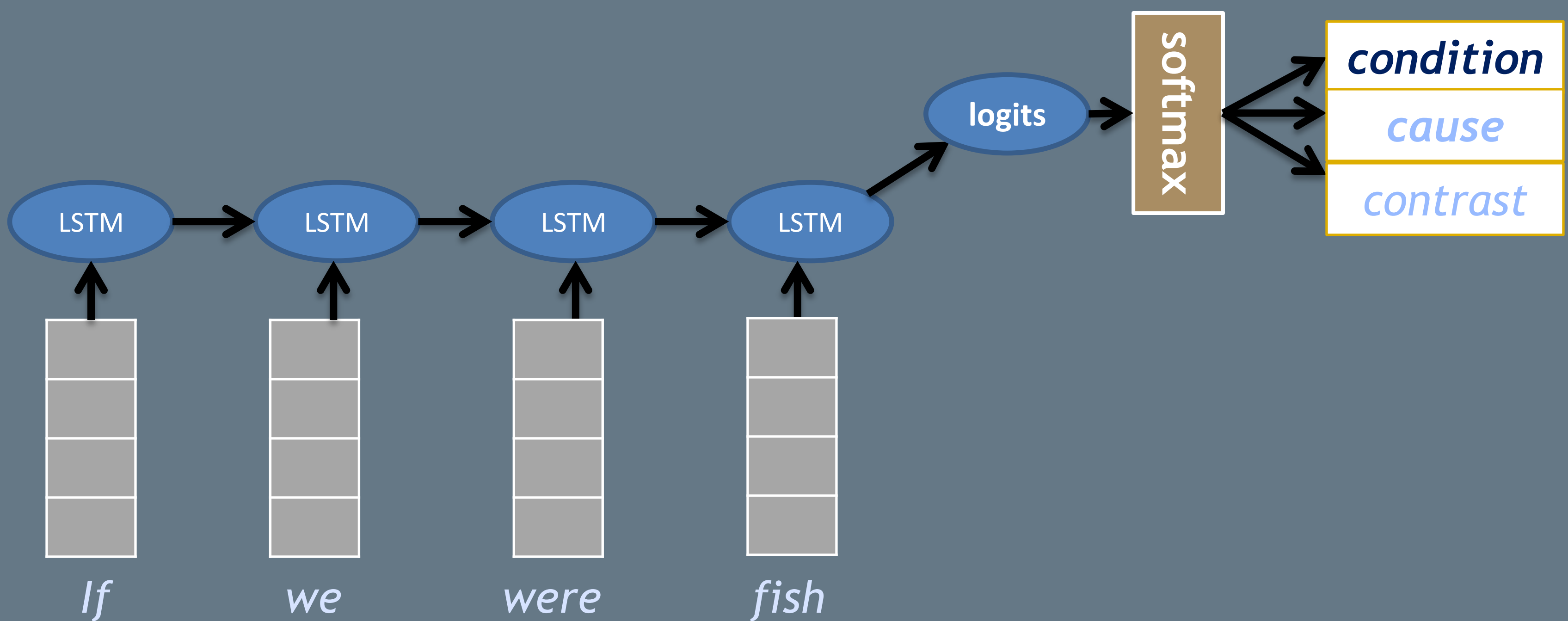
Where's "and"? "but"?

Can we learn contexts?

	most distinctive collexemes by ratio	
relation	**f > 0**	**f > 10**
solutionhood	viable, contributed, 60th, touched, Palestinians	What, ?, Why, did, How
circumstance	holiest, Eventually, fell, Slate, transition	October, When, Saturday, After, Thursday
result	minuscule, rebuilding, distortions, struggle, causing	result, Phoenix, wikiHow, funny, **death**
concession	Until, favoured, hypnotizing, currency	Although, While, though, However, call
justify	payoff, skills, net, Presidential, supporters	NATO, makes, simply, Texas, funny
sequence	Feel, charter, ammonium, evolving, rests	bottles, Place, Then, baking, soil
cause	malfunctioned, jams, benefiting, mandate	because, wanted, religious, projects, **stuff**

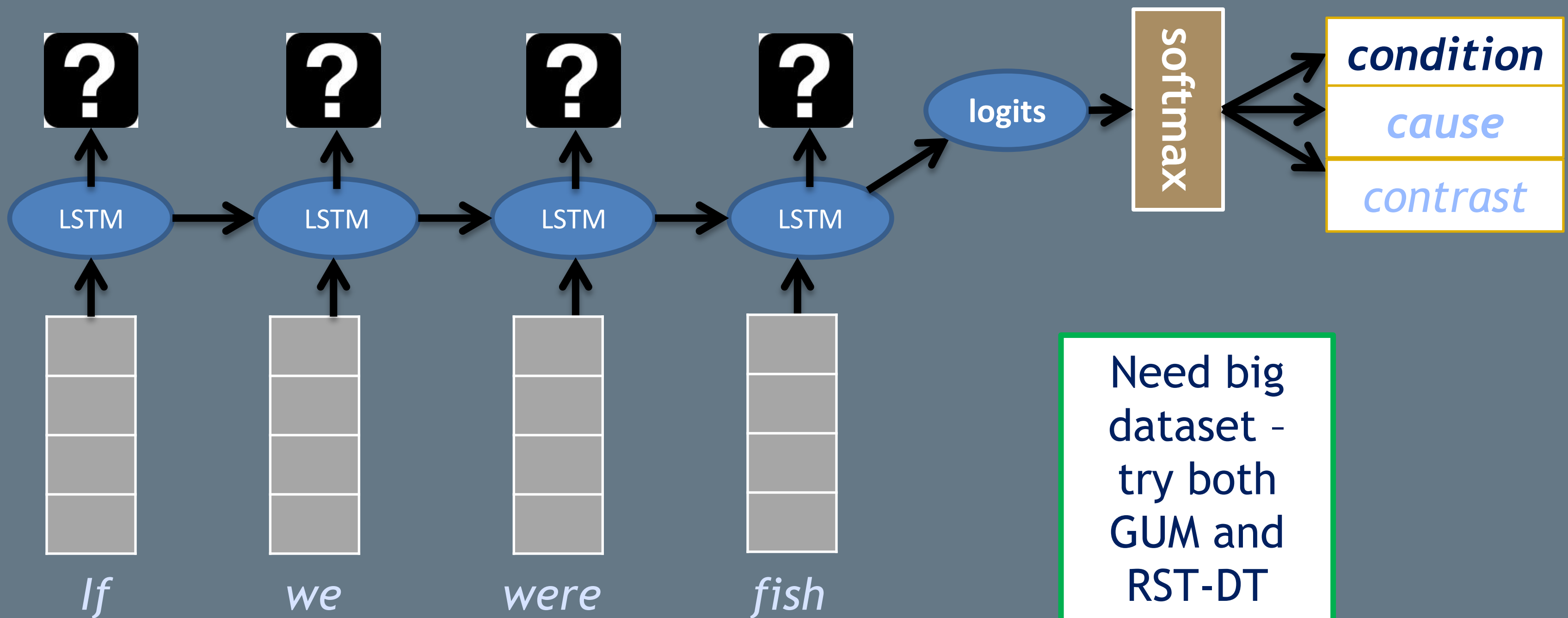
A neural approach with RNNs

- RNNs can recognize relations from text
  (Braud et al. 2017; cf. entailment work, Rocktäschel et al. 2016)
- Can use encoder architecture, single output



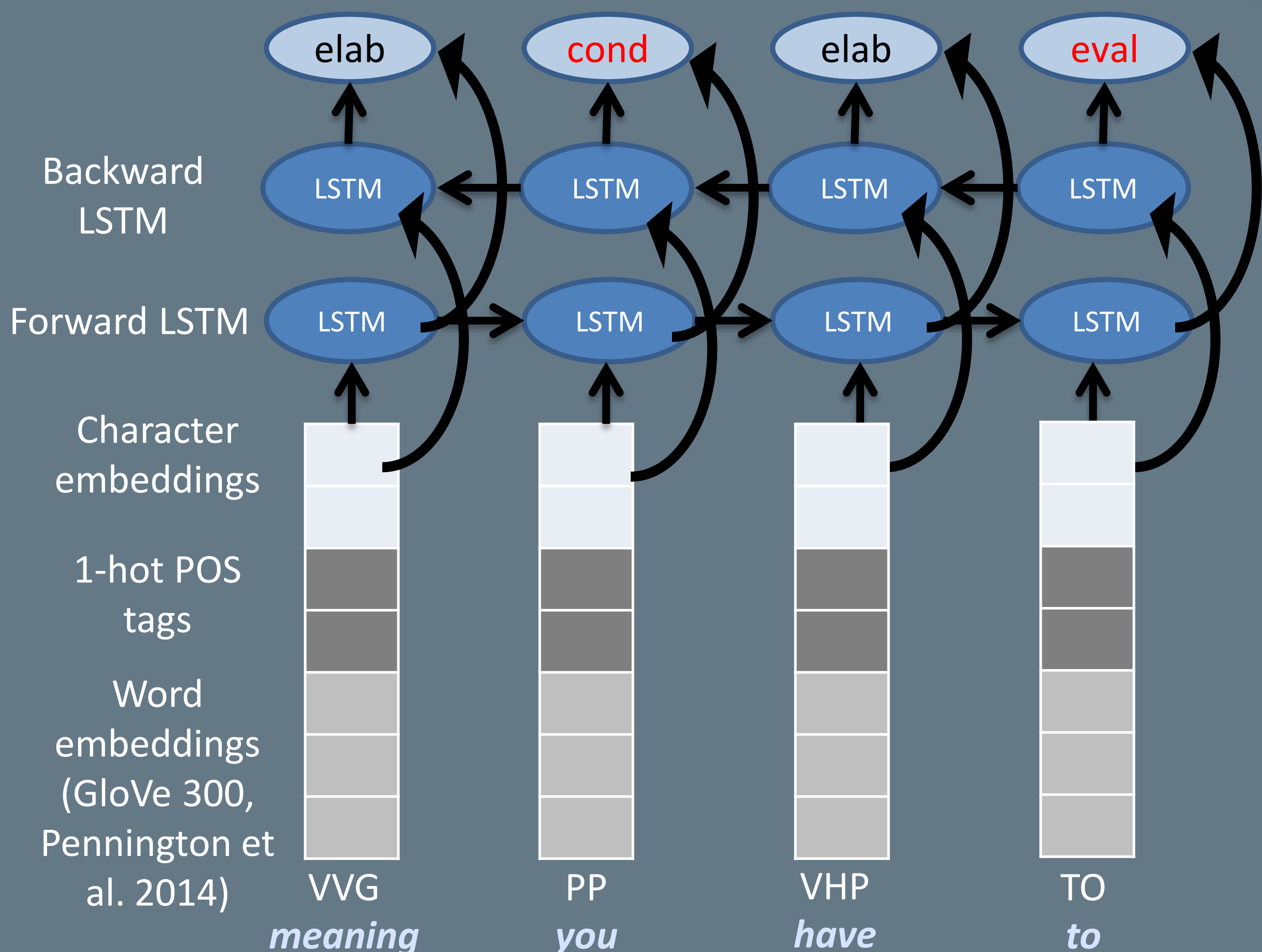
A neural approach with RNNs

- But the LSTM probably already had it as ***If***…
- To find signals, we can listen to output at every token (but loss still based on EDU relation)



Need big dataset – try both GUM and RST-DT

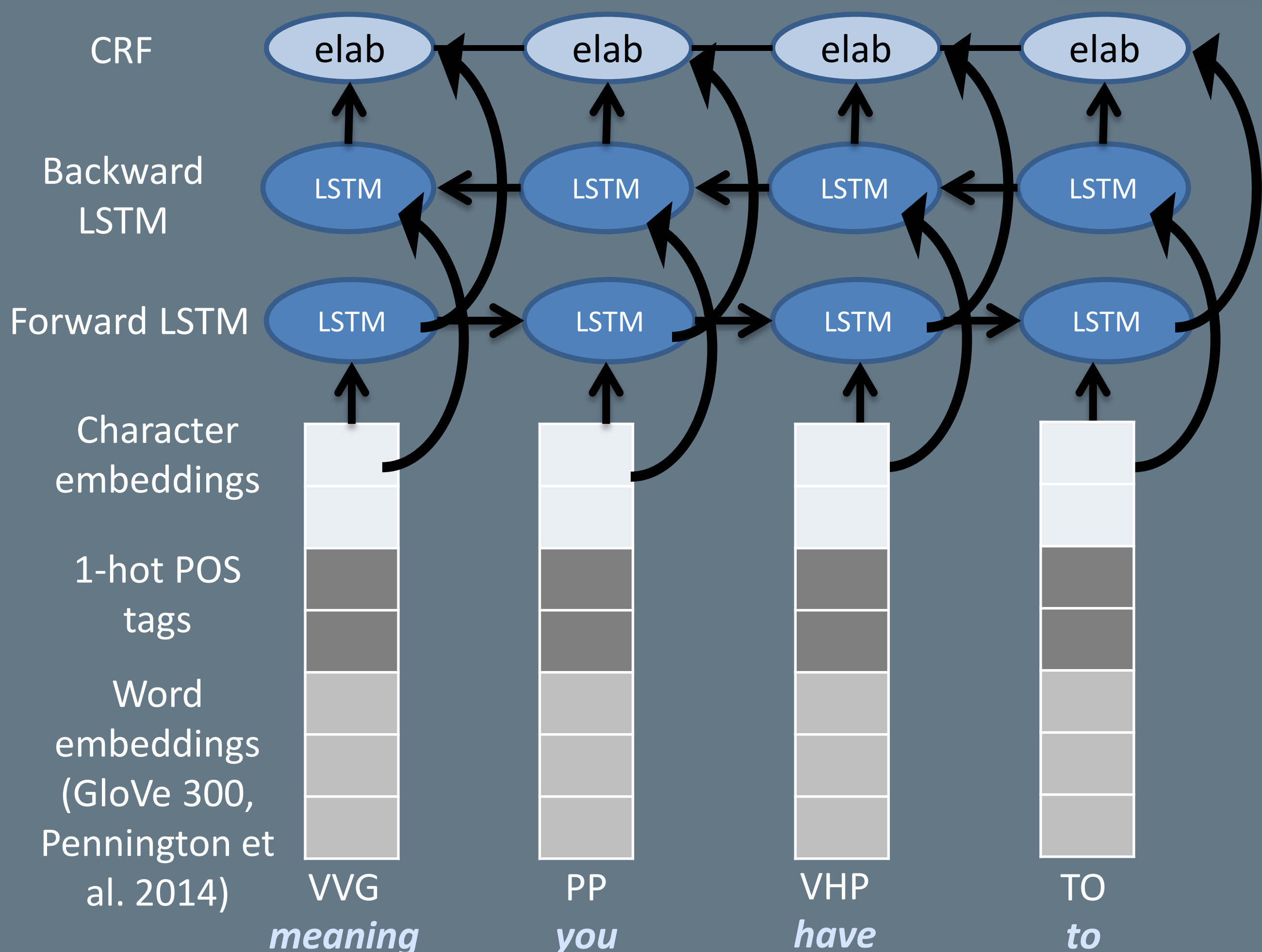
Implemented with BiLSTM (TensorFlow)



Hidden: 300
Optimizer: Adam
(rec.) dropout: 0.5
Minibatch: 20
Activation: tanh

Batch normalization
Trainable embeddings

Adding CRF (Huang et al. 2015, Ma & Hovy 2016)



Hidden:	300
Optimizer:	Adam
(rec.) dropout:	0.5
Minibatch:	20
Activation:	tanh

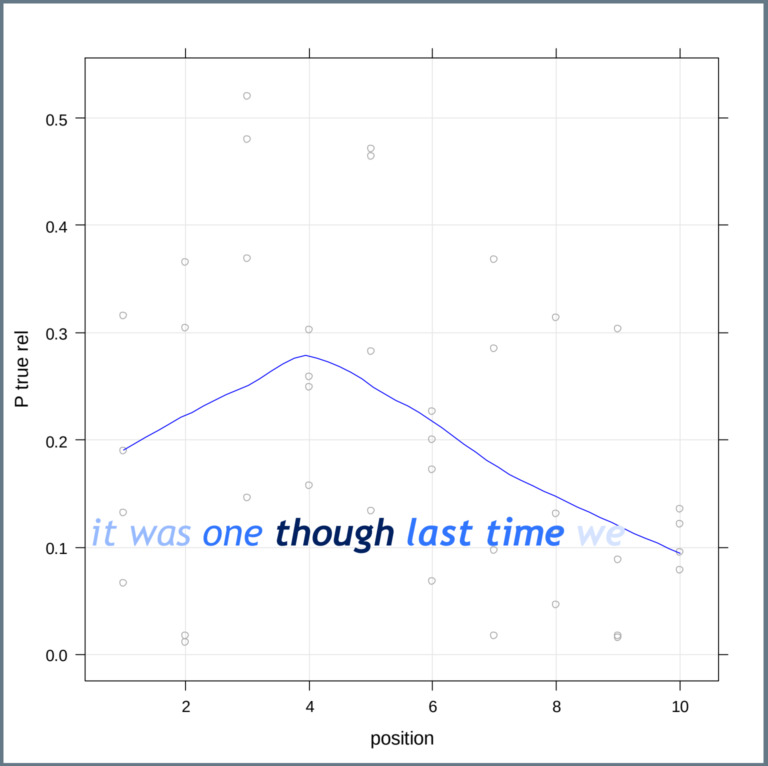
Batch normalization
Trainable embeddings

Single output performance

- Not so interesting, but:
  - RSTDT – relation accuracy by tokens: 47.43% | f1: 41.44
    - Standard train/test split
    - 60 relations [some very rare] – note majority baseline is ~33%
- State of the art on RSTDT, hard to compare:
  - Ji & Eisenstein (2014), using engineered features: 61.75% (by EDUs, 18 relations)
  - Braud et al. (2016), (2017) with RNNs, pretraining on PDTB, coref and more: 60.01% (by EDUs, 18 relations)

Visualizing token-wise softmax

- Basic idea – find the most ‘convincing’ tokens:
  - For each token, output the softmax probability assigned to the correct relation
  - Rank words by probability
  - Shade by average of:
    - Proportion of maximum softmax probability in **sentence**
    - Proportion of maximum softmax probability in **document**



Visualizing token-wise softmax

- [**This** occurs for two reasons **:**]$_{preparation}$ [**As** it moves over land ,]$_{circumstance}$ [it is cut off from the source of energy driving the storm …]$_{cause}$
- [**Combine** 50 milliliters of hydrogen peroxide and a liter of distilled water in a **mixing** bowl .]$_{sequence}$ [A ceramic bowl **will** work **best** ,]$_{elaboration}$ [**but** plastic works **too** .]$_{concession}$

GUM data

Visualizing token-wise softmax

Genre specific knowledge? (GUM)

- [Thursday , May **7** , **2015**]$_{circumstance}$ [The current flag of **New Zealand** .]$_{preparation}$

Word and character embeddings?

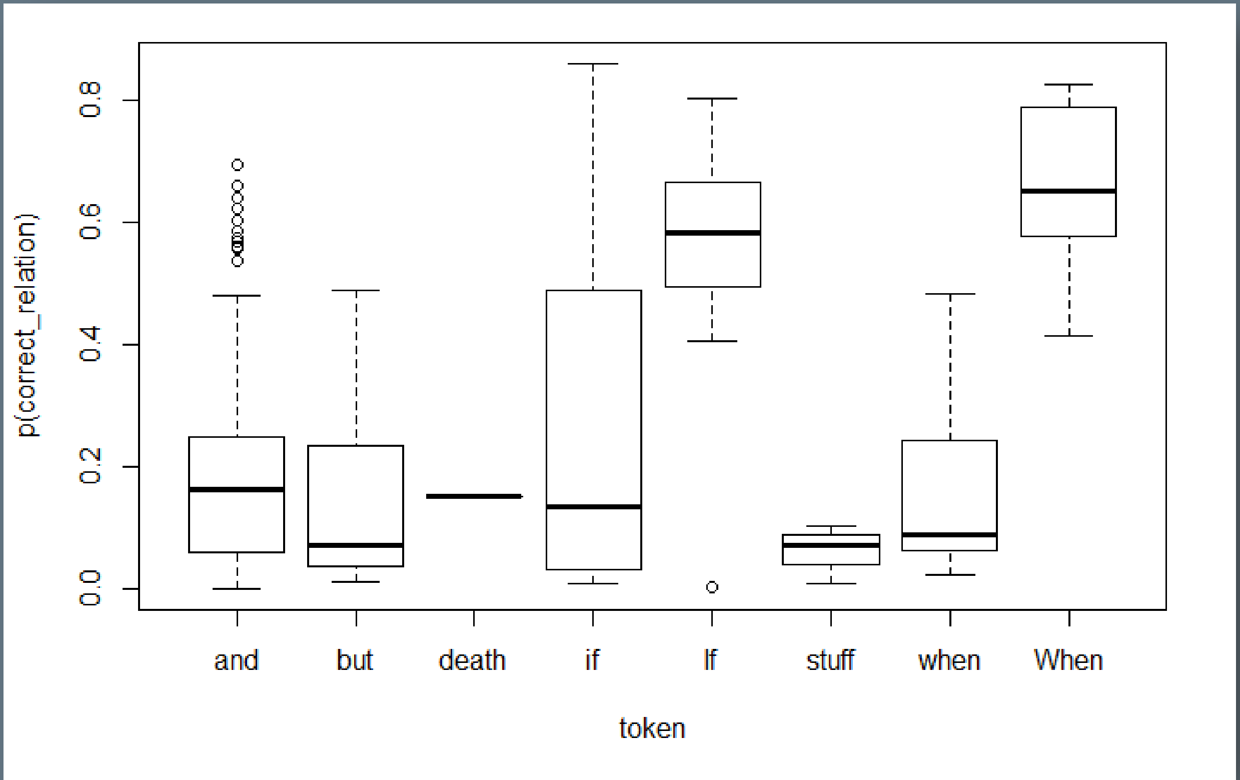
- [I cannot comment **directly** on **how** the Indian government was prepared for this cyclone .]$_{concession}$ [**However ,** the news reports (BBC etc.) were very encouraging]$_{joint}$

Addressing ambiguity

- Reliability of cue words is a big concern:
  - Which cues can we trust?
  - Which cues are we missing because of weak association?

Addressing ambiguity

- We can get ambiguity scores based on range of softmax probabilities (data: GUM)



Addressing ambiguity

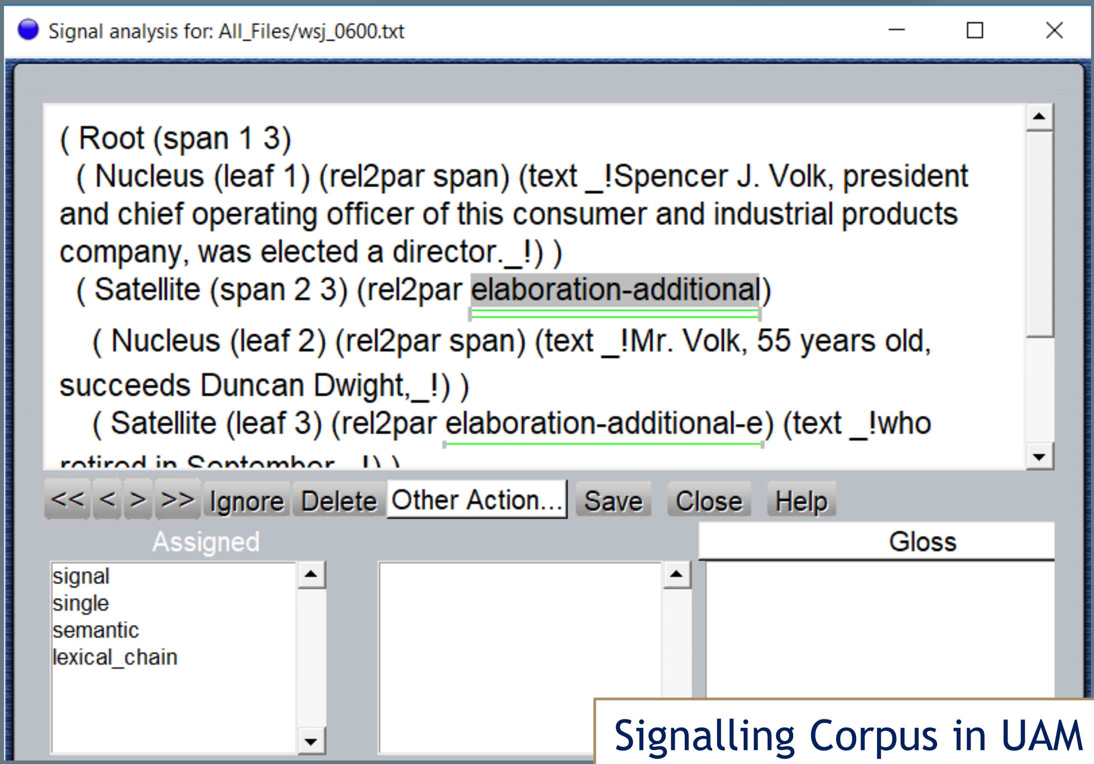
- **Irrelevant 'and's: (RST-DT)**
  - [**but** will continue as a director and chairman of the executive committee .]$_{\text{elaboration}}$
  - [and one began trading on the Nasdaq/National Market System last **week** .]$_{\text{inverted}}$
- **Important 'and's: (RST-DT)**
  - [**and** is involved in claims adjustments for insurance companies .]$_{\text{List}}$
  - [-- **and** from state and local taxes too , for in-state investors .]$_{\text{elaboration}}$

Evaluating signals

- There results are qualitative, non-systematic
- Ideal scenario - compare to ‘gold standard’
  - Use RST-DT Signalling Corpus (Taboada & Das 2013)
  - Open ended annotation of any kind of relation signal:
    - Discourse markers, other expressions
    - Syntactic devices, cohesion
    - Genre conventions…

Evaluating signals

- Problems:
  - Signals annotated at **node** level
  - Non trivial to associate with specific EDUs
  - Location of signal in words is not specified



Signalling Corpus in UAM (O'Donnell 2008)

Toy evaluation

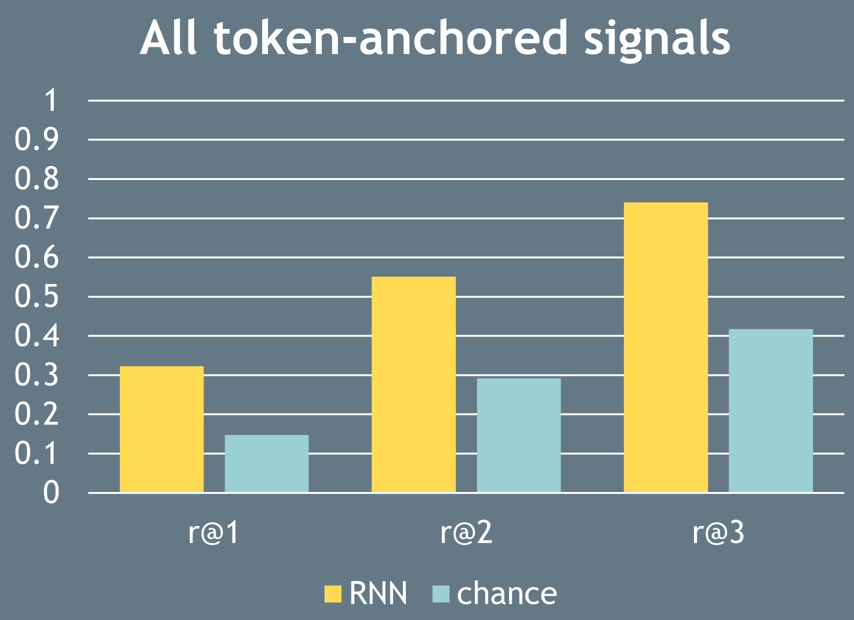
- 3 documents from Signalling Corpus (RST-DT/test)
  - 113 EDUs
  - 210 nodes
  - 153 signals manually inspected
    - Only 83 attributable to a/some tokens
      (not, e.g.: genre, zero relative…)

      In a remark [*someone should remember this time next year,*]
    - Only 47 reasonably detectable by net
      (not, e.g.: lexical chain, syntactic parallelism)

      Congress gave Senator Byrd's state … [***Senator Byrd*** *is chairman..*]

Results

- Network ranks all words (low precision if 0 signals)
- Use ***recall rate @k*** to evaluate

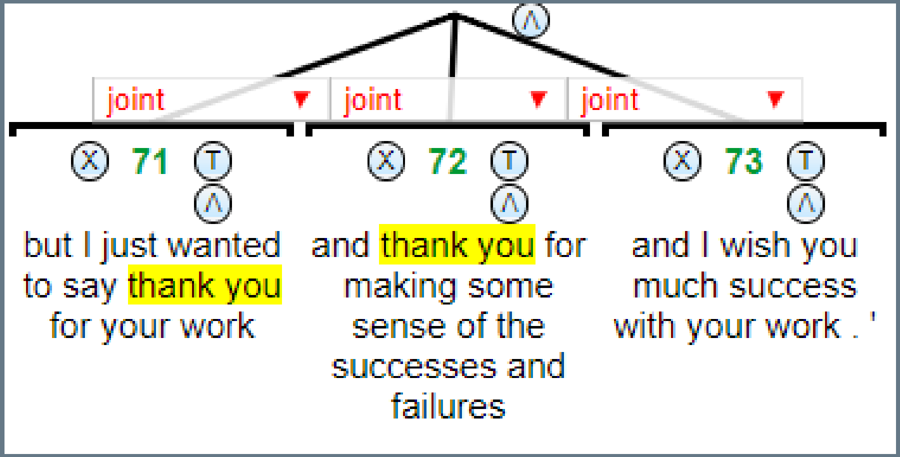


Caveats and WIP

- Network not trained on gold standard (training on relations, not ‘being a signal’)
- Do we want supervised learning on signals?
- Other questions:
  - Can we compare signals across corpora and genres?
  - Are some signals more robust than others?
  - Genre-specific signals?
- Consequences for learning approach?
  ([*Spencer Volk …*][*Mr. Volk…*]$_{\text{elaboration}}$)

Conclusion

- Good times to be working on discourse!
- Multilayer data can expose complex interdependencies
- Some old ideas are now more feasible:
  - From Veins Theory to quantitative DRAs
  - From signal co-occurrence statistics to contextualized LSTM outputs
- We still need new data, new features and new learning approaches!


joint
joint
joint
71
72
73
but I just wanted to say thank you for your work
and thank you for making some sense of the successes and failures
and I wish you much success with your work . '

Thanks!

References

Asher, N./Lascarides, A. 1998. Bridging. *Journal of Semantics,* 15(1), 83–113.
Asher, N./Lascarides, A. 2003. *Logics of Conversation*. Cambridge: Cambridge University Press.
Braud, C./Coavoux, M./Søgaard, A. 2017. Cross-lingual RST discourse parsing. In *Proceedings of EACL 2017*. Valencia, Spain, 292–304.
Braud, C./Plank, B./Søgaard, A. 2016. Multi-view and multi-task training of RST discourse parsers. In *COLING 2016*. Osaka, 1903–1913.
Chiarcos, C./Krasavina, O. 2008. Rhetorical distance revisited: A parametrized approach. In Benz, A./Kühnlein, P. (eds.) *Constraints in Discourse*. Amsterdam/Philadelphia: John Benjamins, 97–115.
Cristea, D./Ide, N./Marcu, D./Tablan, V. 1999. Discourse structure and co-reference: An empirical study. In *Proceedings of the Workshop on the Relationship Between Discourse/Dialogue Structure and Reference*. College Park, MD, 46–53.
Cristea, D./Ide, N./Romary, L. 1998. Veins theory: A model of global discourse cohesion and coherence. In *Proceedings of ACL/COLING*. Montreal, Canada, 281–285.
Geurts, P./Ernst, D./Wehenkel, L. 2006. Extremely randomized trees. *Machine Learning,* 63(1), 3–42.
Hayashi, K./Hirao, T./Nagata, M. 2016. Empirical comparison of dependency conversions for RST discourse trees. In *Proceedings of the SIGDIAL 2016 Conference*. Los Angeles, CA, 128–136.
Hearst, M. A. 1997. TextTiling: Segmenting text into multi-paragraph subtopic passages. *Computational Linguistics,* 23(1), 33–64.
Hirao, T./Yoshida, Y./Nishino, M./Yasuda, N./Nagata, M. 2013. Single-document summarization as a tree knapsack problem. In *Proceedings of EMNLP 2013*. Seattle, WA, 1515–1520.
Huang, Z./Xu, W./Yu, K. 2015. *Bidirectional LSTM-CRF Models for Sequence Tagging*.
Ji, Y./Eisenstein, J. 2014. Representation learning for text-level discourse parsing. In *Proceedings of ACL 2014*. Baltimore, MD, 13–24.
Knott, A. 1996. *A Data-Driven Methodology for Motivating a Set of Coherence Relations*. PhD Thesis, University of Edinburgh.
Knott, A./Dale, R. 1994. Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes,* 18(1), 35–52.
Knott, A./Sanders, T. 1998. The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics,* 30(2), 135–175.
Li, S./Wang, L./Cao, Z./Li, W. 2014. Text-level discourse dependency parsing. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014)*. Baltimore, MD, 25–35.
Ma, X./Hovy, E. 2016. End-to-end sequence labeling via bi-directional LSTM-CNNs-CRF. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL 2016)*. , 1064–1074.
Mann, W. C./Thompson, S. A. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text,* 8(3), 243–281.
O'Donnell, M. 2008. The UAM CorpusTool: Software for corpus annotation and exploration. In *Proceedings of AESLA*. Almeria, Spain, 3–5.
Poesio, . M. Massimo/Vieira, . R. Renata 1998. A corpus-based investigation of definite description use. *Computational Linguistics,* 24(2), 183-216.

References

Polanyi, L. 1988. A formal model of the structure of discourse. *Journal of Pragmatics,* 12(5-6), 601–638.
Prasad, R./Dinesh, N./Lee, A./Miltsakaki, E./Robaldo, L./Joshi, A./Webber, B. 2008. The Penn discourse treebank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*. Marrakesh, Morocco.
Recasens, M./de Marneffe, M.-C./Potts, C. 2013. The life and death of discourse entities: Identifying singleton mentions. In *Proceedings of NAACL 2013*. Atlanta, GA, 627–633.
Roberts, C. 2012. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics,* 5(6), 1–69.
Rocktäschel, T./Grefenstette, E./Hermann, K. M./Kočiský, T./Blunsom, P. 2016. Reasoning about entailment with neural attention. In *International Conference on Learning Representations (ICLR 2016)*.
Sanders, T. J. M./Spooren, W. P./Noordman, L. G. 1992. Towards a taxonomy of coherence relations. *Discourse Processes,* 15, 1–35.
Stede, M./Grishina, Y. 2016. Anaphoricity in connectives: A case study on German. In *Proceedings of the Workshop on Coreference Resolution Beyond OntoNotes (CORBON 2016)*. San Diego, CA, 41–46.
Sweetser, E. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: CUP.
Taboada, M./Das, D. 2013. Annotation upon annotation: Adding signalling information to a corpus of discourse relations. *Dialogue and Discourse,* 4(2), 249–281.
Taboada, M./Lavid, J. 2003. Rhetorical and thematic patterns in scheduling dialogues: A generic characterization. *Functions of Language,* 10(2), 147–179.
Tetreault, J./Allen, J. 2003. An empirical evaluation of pronoun resolution and clausal structure. In *Proceedings of the 2003 International Symposium on Reference Resolution and its Applications to Question Answering and Summarization*. Venice, 1–8.
Toldova, S./Pisarevskaya, D./Ananyeva, M./Kobozeva, M./Nasedkin, A./Nikiforova, S./Pavlova, I./Shelepov, A. 2017. Rhetorical relation markers in Russian RST treebank. In *Proceedings of the 6th RST Workshop*. Santiago de Compostela, Spain, 29–33.
Vieira, . R. Renata/Poesio, . M. Massimo 2000. Corpus-based development and evaluation of a system for processing definite descriptions. In *Proceedings of the 18th conference on Computational linguistics*. Stroudsburg, PA: Association for Computational Linguistics, 899–903.
van der Vliet, N./Redeker, G. 2014. Explicit and implicit coherence relations in Dutch texts. In Gruber, H./Redeker, G. (eds.) *The Pragmatics of Discourse Coherence: Theories and applications*. Amsterdam/Philadelphia: John Benjamins, 23–52.
Webber, B. L. 1988. *Discourse Deixis: Reference to Discourse Segments*. Technical Report, University of Pennsylvania.
Zeldes, A. 2017a. A distributional view of discourse encapsulation: Multifactorial prediction of coreference density in RST. In *6th Workshop on Recent Advances in RST and Related Formalisms*. Santiago de Compostela, Spain.
Zeldes, A. 2017b. The GUM corpus: Creating multilayer resources in the classroom. *Language Resources and Evaluation,* 51(3), 581–612.
Zeldes, A./Simonson, D. 2016. Different flavors of GUM: Evaluating genre and sentence type effects on multilayer corpus annotation quality. In *Proceedings of LAW X – The 10th Linguistic Annotation Workshop*. Berlin, 68–78.

Sentence type annotation

- Extended version of SPAAC scheme (Leech et al. 2003; not created for this study)

tag	type	example
q	polar yes/no question	*Did she see it?*
wh	WH question	*What did you see?*
decl	declarative (indicative)	*He was there.*
imp	imperative	*Do it!*
sub	subjunctive (incl. modals)	*I could go*
inf	infinitival	*How to Dance. Why not go?*
ger	gerund-headed clause	*Finding Nemo. Hiring employees*
intj	interjection	*Hello!*
frag	fragment	*The End.*
other	other predication or combination	*Nice, that!* Or: *'I've had it, go!'* (decl+imp)

Only weak correlations…

- For all EDU pairs:
  - Most have 0 coreference
  - Especially direct antecedents have very low distance
  - Not much predictability (cf. Tetreault & Allen) ✓ ☹