

# disentangling chat with local coherence models

micha elsner and eugene charniak

bitbucket.org/melsner/browncoherence

## chat disentanglement

problem: cluster chat utterances into threads

each conversation is **coherent**  
utterances follow from context  
but the unthreaded transcript isn't  
...so search for coherent threads

Why doesn't Linux find a dhcp server?

I don't eat bananas.

The fact that there isn't one?

You should, they have potassium goodness.

I eat cardboard boxes because of the fiber.

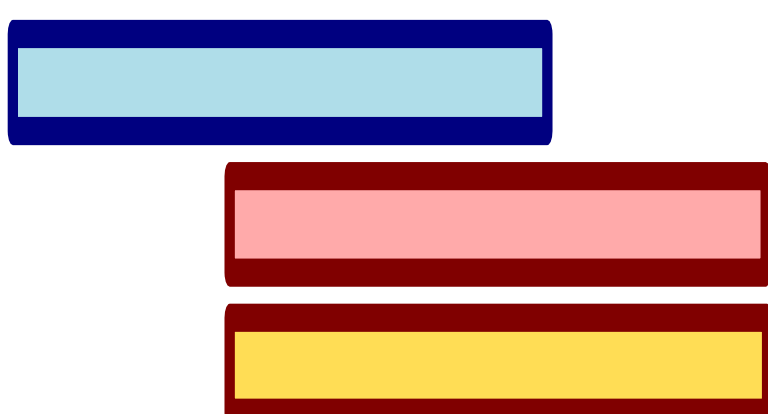
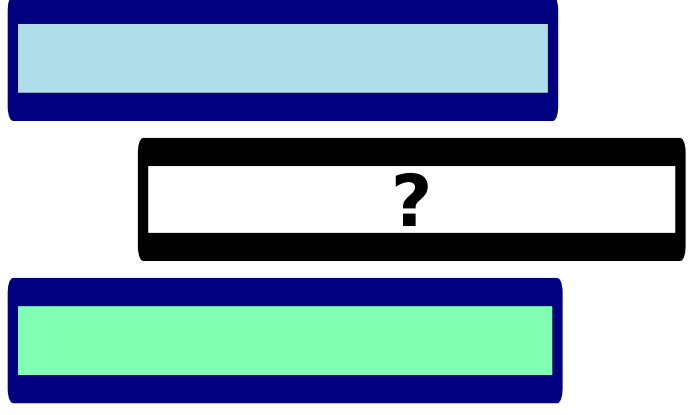
## do popular models of coherence work on this task?

despite being developed for news text summarization?

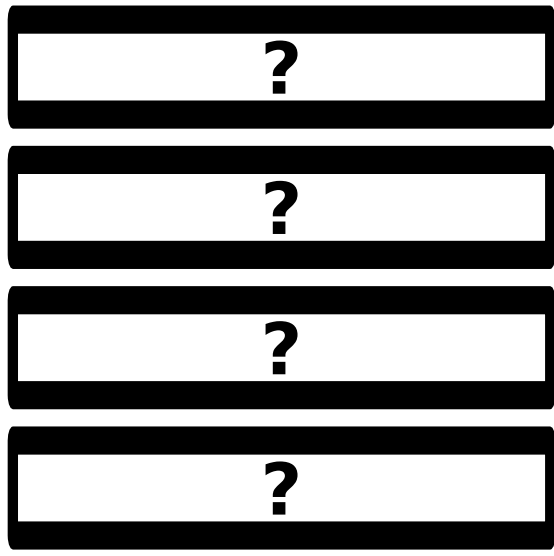
## coherence models

- entity grid noun repetitions and syntactic roles [Lapata+Barzilay '05]
- topical entity grid egrid using LDA similarity instead of repetitions
- IBM-1 'translates' prev sentence to next [Soricut+Marcu '06]
- pronouns checks if pronouns can find antecedents [C+E '09]
- chat-specific utterance time, speaker, name mentions [E+C '08]
- combined discriminatively trained log-linear mixture
- EC'08 [Elsner+Charniak '08]  
baseline with chat-specific + simple lexical features

## single-utterance disentanglement

		swbd	IRC
gold clusters		chat-specific 58	74
		egrid 77	+chat-specific 79
		topical 78	77
one unknown		IBM 69	76
		pronouns 52	74
algorithmically easy!		combined 83	78
		EC'08 76	76

## full-scale disentanglement

		swbd	IRC
all unknown		chat-specific 55	results poor:
		egrid 59	coherence alone
		topical 60	creates too many
		IBM 56	threads
tabu search for clustering		pronouns 54	
one-to-one overlap score		combined 64	future work!
		EC'08 58	

## conclusions

most models work on disentanglement  
...but domain matters for lexical models  
unlexicalized entity grid often best  
...domain insensitive, multi-sentence context

### coherence modeling research

disentanglement as evaluation task  
can measure domain dependence

### chat research

better topic models, pronoun coref for chat