

Lili Kotlerman, Ido Dagan
Bar Ilan University

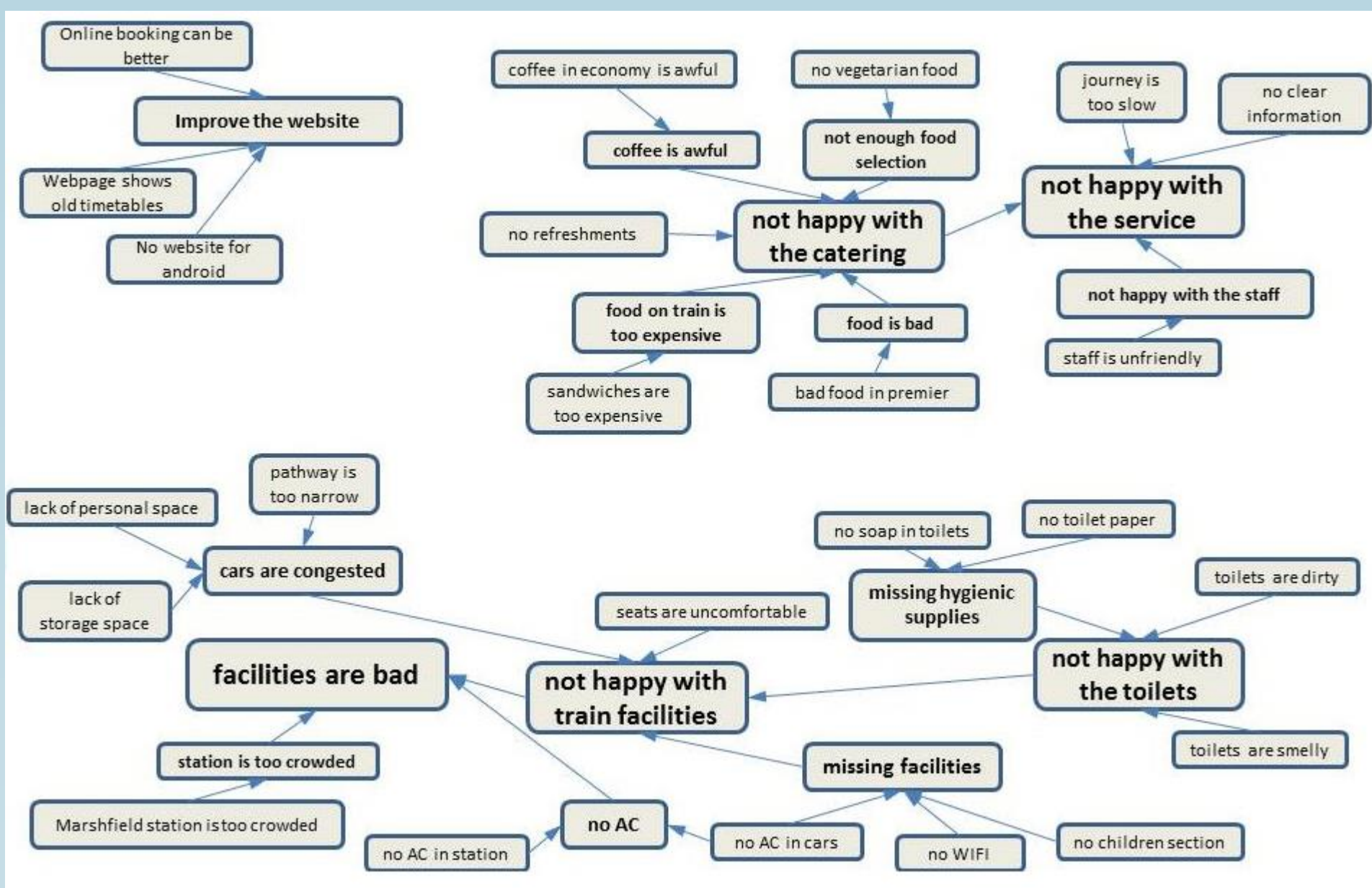
Bernardo Magnini, Luisa Bentivogli
Fondazione Bruno Kessler

Abstract

We introduce a novel type of graphs for NLP – Textual Entailment Graphs (TEG). We present the first gold-standard dataset of TEGs, describe a methodology we developed for their construction and evaluate state-of-the-art RTE engines for this task. We show the utility of TEGs for text exploration and suggest that TEGs might have different usages and that their automatic creation is an interesting task for the community.

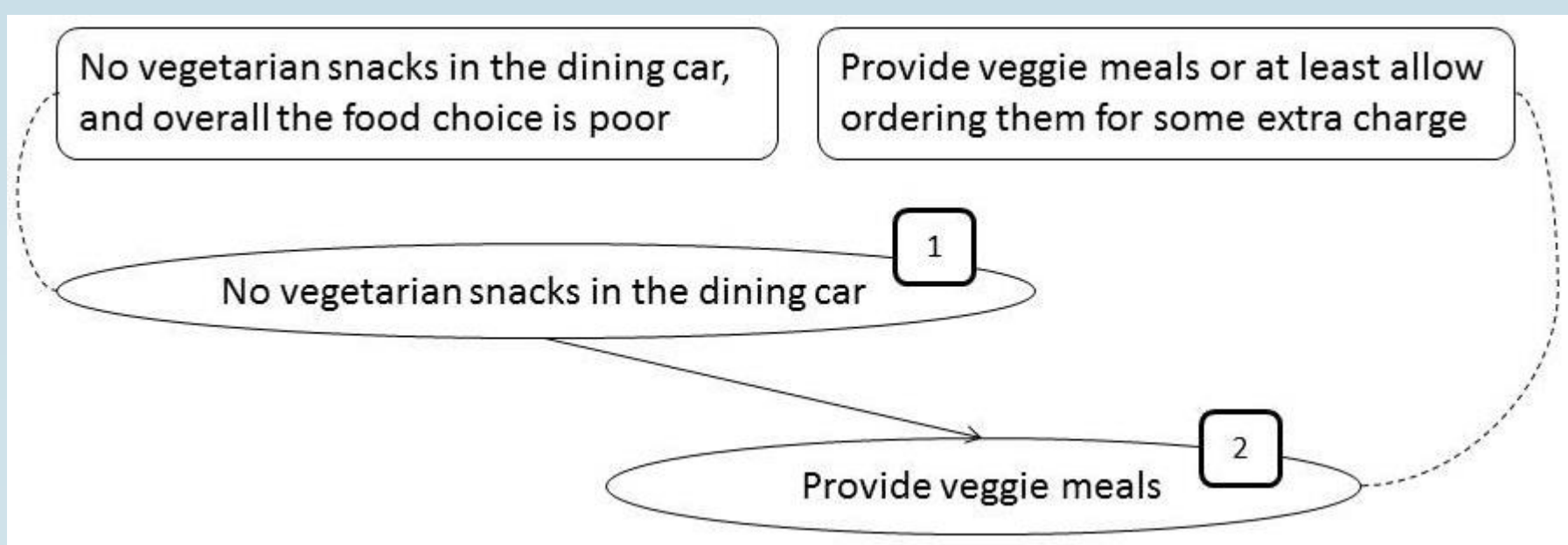
Textual Entailment Graph (TEG)

- ✓ Nodes: textual fragments (content unit conveying a single complete proposition and made up of a sequence of words from an original text.)
- ✓ Edges: entailment relations
- ✓ Transitivity constraint

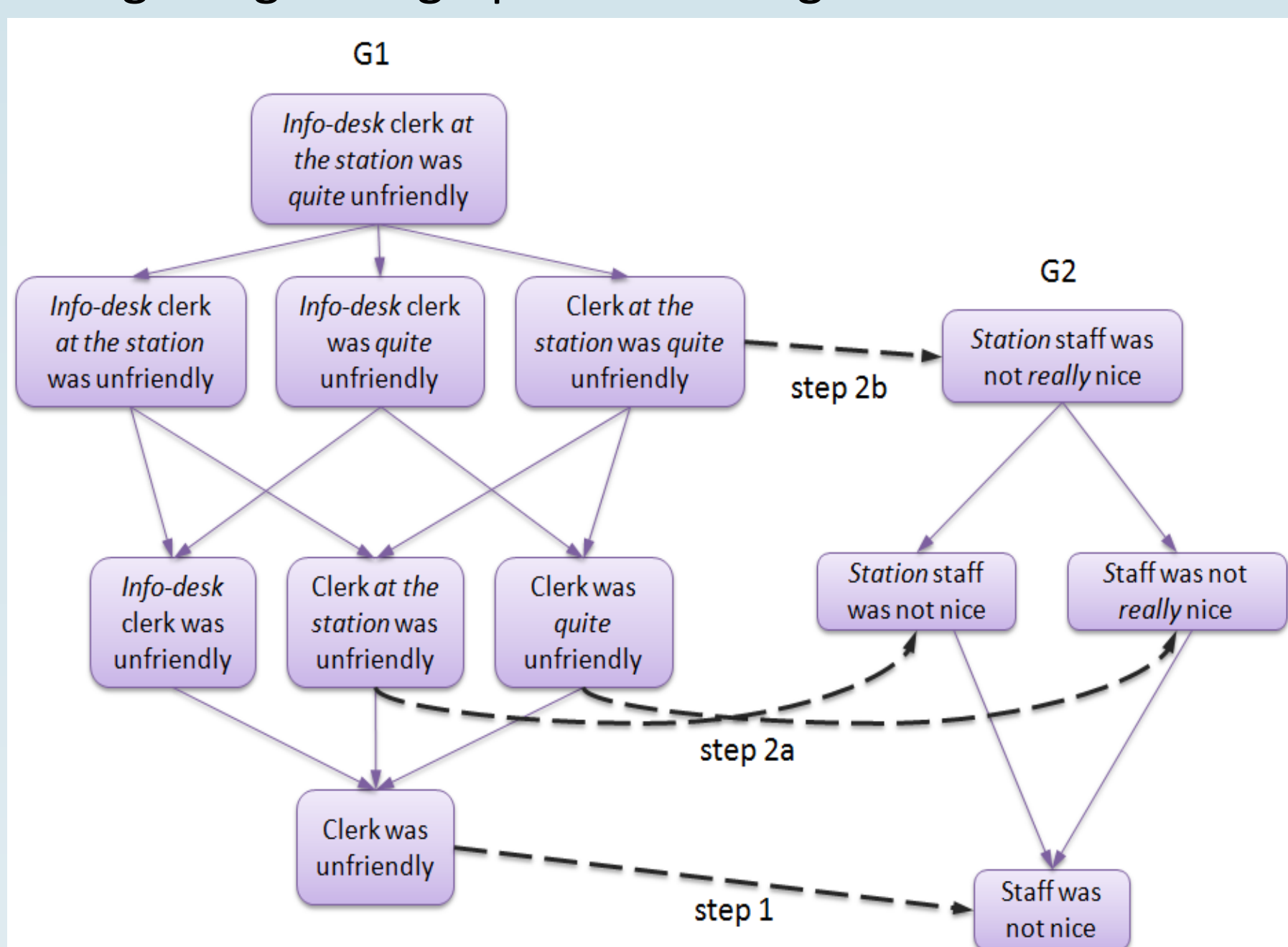


TEG Generation Methodology

- ✓ Extract fragments of interest from the input set of texts



- ✓ Construct fragment entailment graphs
- ✓ Merge fragment graphs into a single TEG



Baselines for Automatic TEG Construction

Manually annotated dataset. Source: over 200 user e-mails

	Clusters	Original texts	Fragment Graphs	Total nodes	Total edges	Intra-fragment edges	Inter-fragment edges
Dev Set	14	202	233	360	3555	160	3395
Test Set	7	64	79	155	2001	120	1881

State-of-the-art RTE engines: Excitement Open Platform

Merge procedure	TE engine	Initial			Pairwise comparisons
		R	P	F1	
All pairs	TIE	1.00	0.29	0.40	14,036
	TIE+WN+VO	0.45	0.45	0.38	
	BIUTE+WN+VO	0.96	0.34	0.45	
Structure-based	TIE	0.52	0.34	0.35	6,488
	TIE+WN+VO	0.45	0.35	0.35	6,066
	BIUTE+WN+VO	0.49	0.39	0.36	5,734

Merge procedure	TE engine	Clique		
		R	P	F1
All pairs	TIE	1.00	0.30	0.42
	TIE+WN+VO	1.00	0.30	0.43
	BIUTE+WN+VO	1.00	0.35	0.47
Structure-based	TIE	0.89	0.39	0.49
	TIE+WN+VO	0.90	0.43	0.54
	BIUTE+WN+VO	0.91	0.44	0.55

TEGs for Text Exploration

- ✓ Generalization
- ✓ Browsing and drill-down
- ✓ Quantitative analysis

