ABCD: A Graph Framework to Convert Complex Sentences to a Covering Set of Simple Sentences

Yanjun Gao, Ting-hao Huang, Rebecca Jane Passonneau
{yug125, txh710, rjp49} @ psu.edu
Pennsylvania State University
ACL 2021

nlp penn state university

Crowd-AI Lab
crowdailab.net

Introduction

Orig	Sokuhi was born in Fujian and was ordained at 17.
SS1	Sokuhi was born in Fujian.
SS2	Sokuhi was ordained at 17.

- ❖ One of many sophisticated complex sentence types we are dealing with!
- ❖ Others include: embedded clauses, relative clauses, etc.

Introduction

Orig	Sokuhi was born in Fujian and was ordained at 17.
SS1	Sokuhi was born in Fujian.
SS2	Sokuhi was ordained at 17.

- ❖ Decomposing complex sentences into simple sentences is a fundamental and significant research in NLP.
- ❖ Propositions are addressed in *summarization, argument mining, question answering, knowledge graph construction* etc.

Previous Automated Methods

- Rule-based parsing
  - Pros: hierarchical processing
  - Cons: less flexibility; limited performance from rules
- Neural text segmentation
  - Pros: easy to train
  - Cons: output is incomplete propositions
- Encoder-Decoder based Split-and-Rephrase
  - Pros: output complete propositions
  - Cons: hard to train; hard to generalize

Previous Automated Methods

- Rule-based parsing
  - Pros: hierarchical processing
  - Cons: less flexibility; limited performance from rules
- Neural text segmentation
  - Pros: easy to train
  - Cons: output is incomplete propositions
- Encoder-Decoder based Split-and-Rephrase
  - Pros: output complete propositions
  - Cons: hard to train; hard to generalize

Our work: ABCD: A linguistic-aware, neural "editor" to learn to edit sentence graphs into subgraphs

ABCD Methodology

copy drop break

Complex Sentence: Sokuhi was born in Fuzhou, China and was ordained at 17 by Feiying Tongrong.

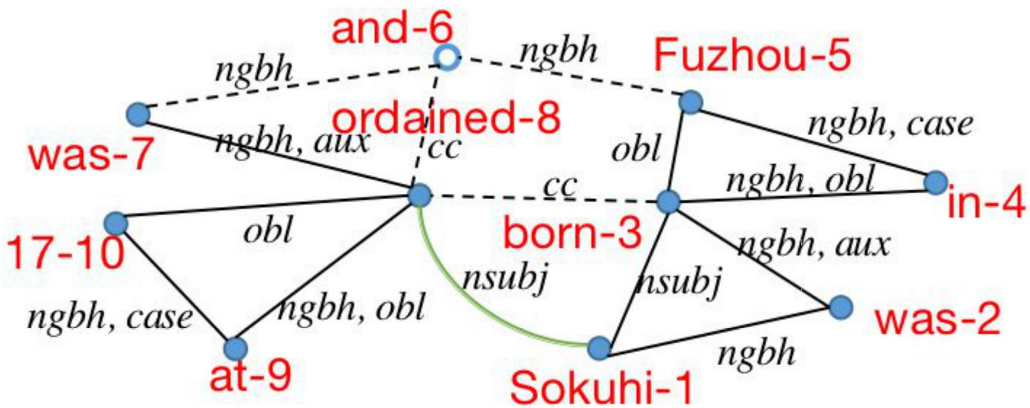
Clauses: 1. Sokuhi was born in Fuzhou.
2. Sokuhi was ordained at 17 by Feiying Tongrong.

- Converts sentences into **word relation graphs** (WRGs) that encode word *adjacency* and *grammatical dependencies* (e.g., subject dependencies)
- **Four graph edge edit** types:
  - **Accept**
  - **Break**
  - **Copy**
  - **Drop**

ABCD Methodology

- **Distant supervision label** creator generating ground truth edge edits
- **Neural model** learns to classify 4 edit types
- **Postprocesser with DFS algorithm**
  - finds connected components (CC)
  - converts CC into sentences

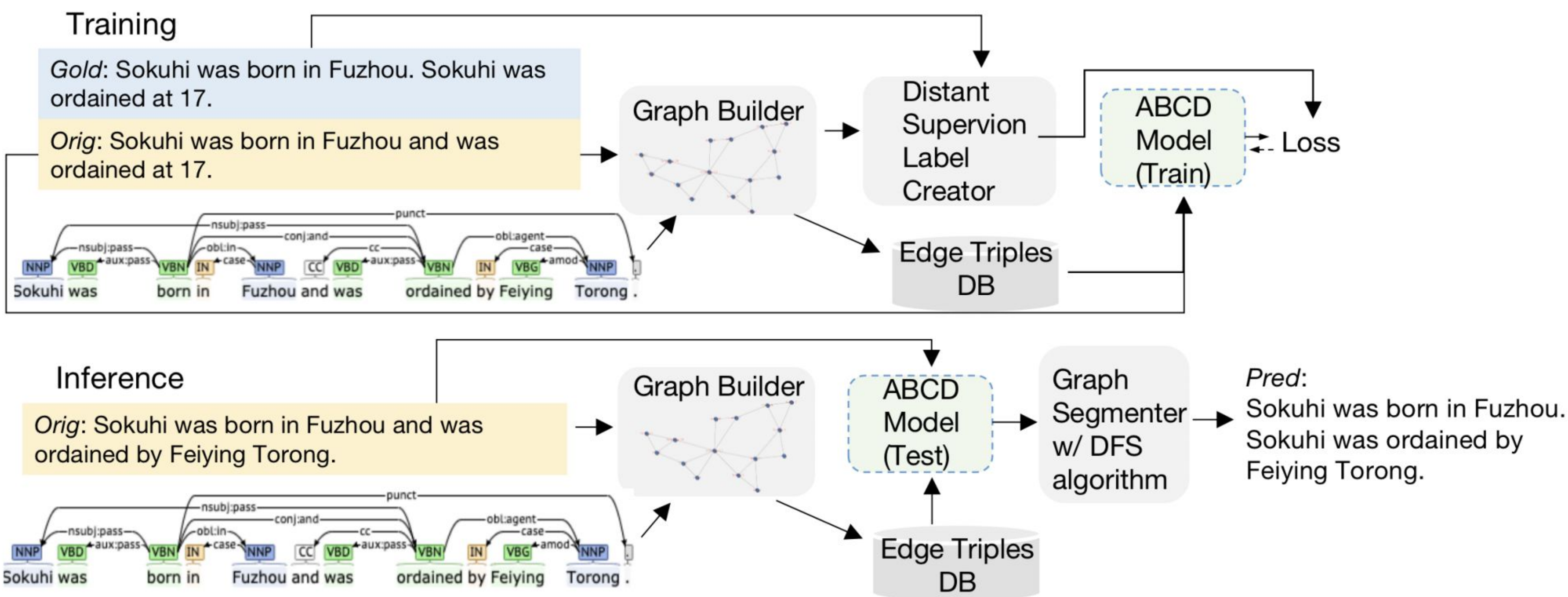
Example WRG



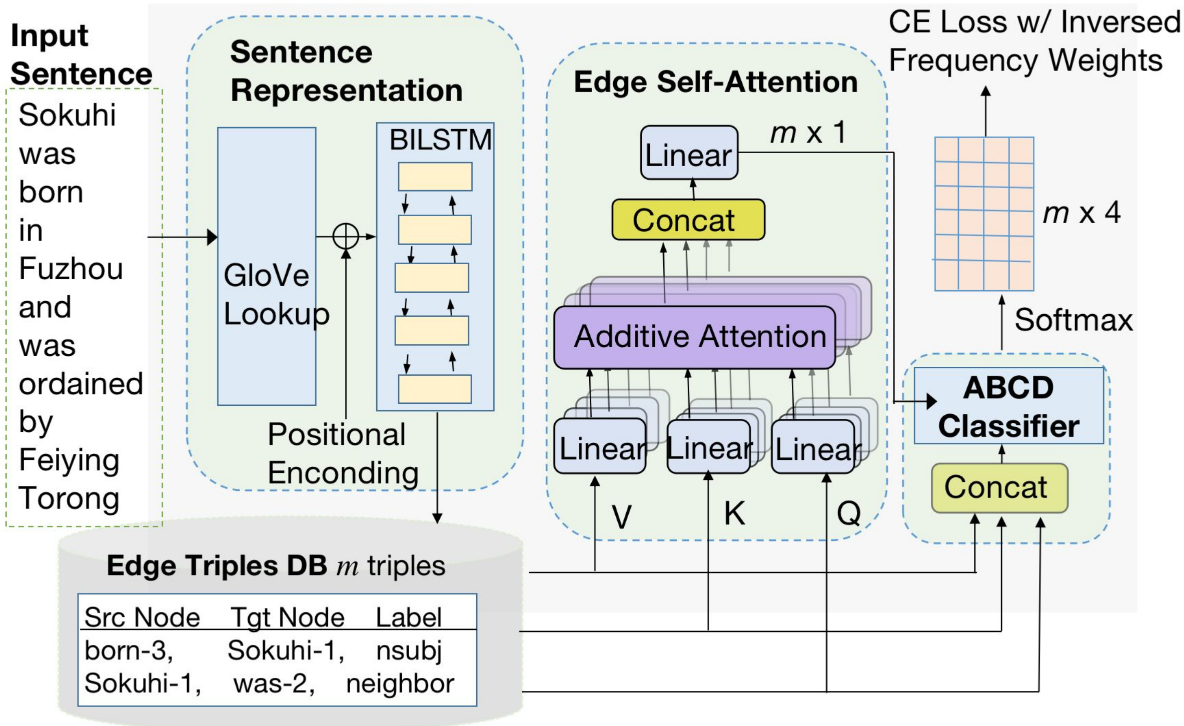
Src Node	*Tgt Node*	*Label(s)*
born-3,	Sokuhi-1,	nsubj
Sokuhi-1,	was-2,	ngbh
was-2,	born-3,	ngbh,aux
...	...	...

Orig	Sokuhi was born in Fuzhou and was ordained at 17.
SS 1	Sokuhi was born in Fuzhou.
SS 2	Sokuhi was ordained at 17.

ABCD Pipeline



ABCD Neural Model



Dataset

- **DeSSE:** Decomposed Students' Essays
- **MinWiki:** from Wikipedia Text

Dataset	A	B	C	D
MinWiki	85.23%	4.58%	3.60%	6.57%
DeSSE	74.77%	2.39%	5.62%	17.21%
MinWiki	0.0167	0.3533	0.4164	0.2135
DeSSE	0.0200	0.6266	0.2658	0.0876

Table 2: Distributions (Top) and inverse class weights (Bottom) for the four edit labels on both MinWiki and DeSSE datasets.

Intrinsic Evaluation on Output Propositions

Group	Model	MinWiki #T /SS	Match #SS(%)	BLEU4	BERTSc
Parsing	DisSim	8.50	68.46	64.20	94.42
	DCP_{vp}	14.82	45.49	28.80	64.50
	DCP_{sbar}	19.07	17.49	19.35	49.07
	DCP_{recur}	16.30	67.90	31.78	58.08
Encoder-decoder	COPY	9.37	**79.26**	**80.96**	**95.96**
ABCD biLSTM	mlp	9.37	78.61	75.80	92.91
	bilin	9.53	76.72	76.38	90.28

Table 4.4. Performance of baselines and our models on Minwiki test set (N=1075, #T/SS = 10.03). We report numbers of token per propositions (#T/SS), number of input sentences that have match number of output between prediction and ground truth in percentage (Match #SS%), BLEU with four-gram and BERTScore.

Intrinsic Evaluation on Output Propositions

Group	Model	DeSSE #T /SS	Match #SS(%)	BLEU4	BERTSc
Parsing	DisSim	9.59	40.00	37.89	89.54
	DCP_{vp}	15.99	42.40	47.25	60.18
	DCP_{sbar}	17.24	44.81	48.02	59.89
	DCP_{recur}	14.16	**55.63**	34.44	61.37
Encoder-decoder	COPY	18.13	36.20	45.91	88.71
ABCD biLSTM	mlp	8.85	53.29	**53.42**	90.23
	bilin	8.10	52.66	41.57	**94.78**

Table 4.5. Performance of baselines and our models on DeSSE test set (N=790, #T/SS =9.07). We report numbers of token per propositions (#T/SS), number of input sentences that have match number of output between prediction and ground truth in percentage (Match #SS%), BLEU with four-gram and BERTScore.

Output Example

Orig: I guess I always knew it was **genetics** but I didn't know why our features are the way that they are.
Human: I guess I always knew it was genetics. I didn't know why our features are the way that they are. (n=2)
Copy: I guess I always knew it was interesting but I didn't know why our features are the way that they are. (n=1)
ABCD: I guess I always knew it was genetics. I didn't know why our features are the way that they are. (n=2)

Future Work

- ❖ Introducing **sequential decision** into the model
- ❖ Testing ABCD **in downstream applications**

ABCD is available at https://github.com/serenayj/ABCD-ACL2021

DeSSE is available at https://github.com/serenayj/DeSSE