NAACL 2021

Semantic Frame Forecast

Chieh-Yang Huang and Ting-Hao Kenneth Huang
Penn State University, PA, USA
{chiehyang, txh710}@psu.edu

Crowd-AI Lab
crowdailab.net

PennState

Introduction

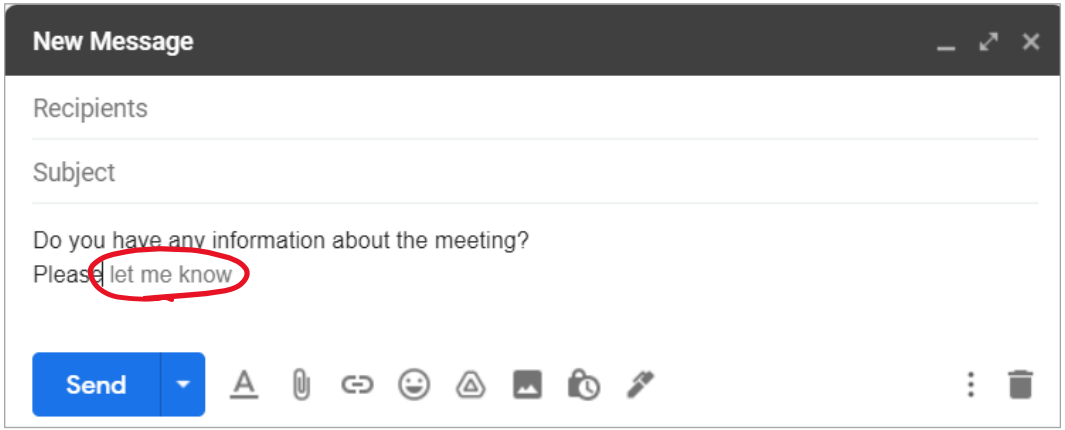
- Story writing is hard. Writers can struggle to develop the follow-up scenes.

  "*Anyone who says writing is easy isn't doing it right.*" - Amy Joy.

➡ Provide hints or texts to help!

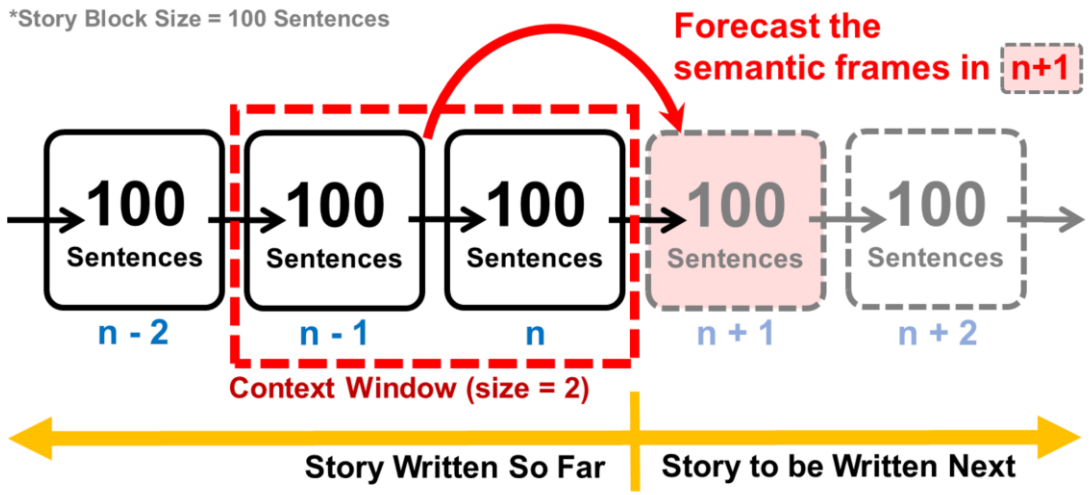
Introduction

- However, most of the tools aim at providing ideas or texts for **the immediate future** (one to a few sentences later).



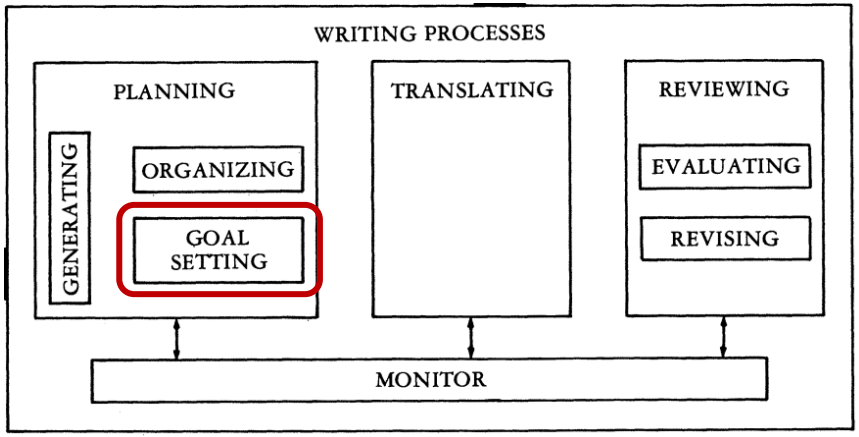
Introduction

- We present a **frame representation** on **story blocks** to handle (1) a longer history and (2) a longer future.



What To Predict?

- Semantic frames are high-level concepts of **things.**
- Forecasting a frame means predicting what would happen next!
- **Goal setting** in the writing process theory [1]!



[1] Flower, Linda, and John R. Hayes. "A cognitive process theory of writing." College composition and communication 32.4 (1981): 365-387.

Semantic Frame Representation

A Story Block (20 Sentences)

His mouth turned up slightly at the corners. “Are you really.”

***

Meanwhile, in the city center…

too small for his enormous hands. There were two huge guns strapped in leather holsters on his hips.

B Semantic Frames

His mouth turned up slightly at the corners. “Are you really.”

***

Meanwhile, [in] the [city] [center]…

[There] [were] [two] [huge] [guns] strapped [in] leather holsters [on] his hips.

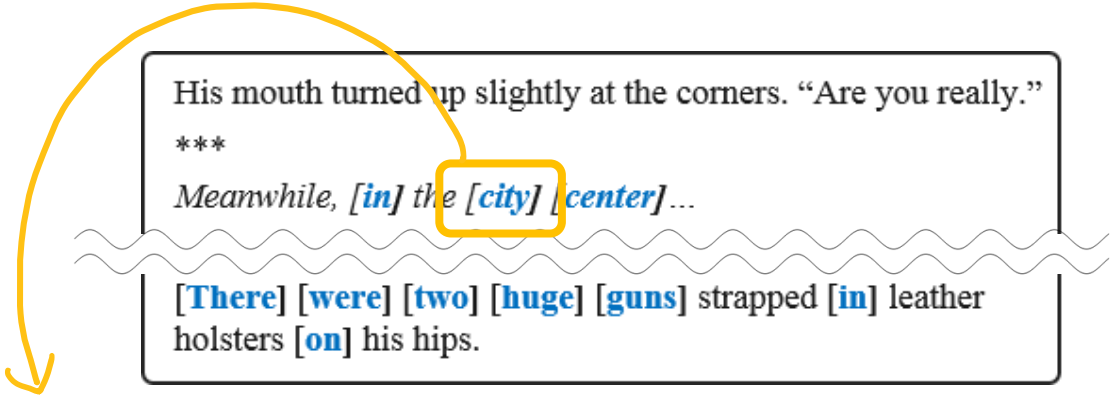
Parse Semantic Frame using Open-Sesame [2]

* **FrameNet** [3] defines a total of 1,221 high-level conceptual units, which are called **frames**.

[2] Baker, Collin F., Charles J. Fillmore, and John B. Lowe. "The berkeley framenet project." 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, Volume 1. 1998.

[3] Swayamdipta, Swabha, et al. "Frame-semantic parsing with softmax-margin segmental rnns and a syntactic scaffold." arXiv preprint arXiv:1706.09528 (2017).

A Semantic Frame



Political_locales

Definition This frame covers words that name Locations as defined politically, or administratively.

Lexical Units
(Words that can trigger the frame.)

*barony.n, borough.n, city-state.n, **city.n**, commonwealth.n, country.n, county.n, diocese.n, district.n, duchy.n, empire.n, federal.a, fiefdom.n, global.a, international.a, internationally.adv, jurisdiction.n, kingdom.n, land.n, local.a, locality.n, megalopolis.n, metropolis.n, multinational.a, municipal.a, municipality.n, nation.n, national.a, parish.n, principality.n, province.n, provincial.a, realm.n, state [internal].n, state [sovereign].n, territorial.a, territory.n, town.n, township.n, village.n, viscountcy.n, world.n*

Semantic Frame Representation

B Semantic Frames

His mouth turned up slightly at the corners. “Are you really.”

***

Meanwhile, [in] the [city] [center]...

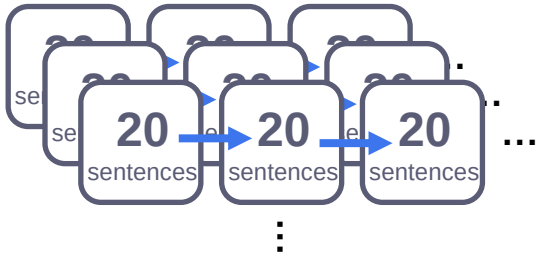
[There] [were] [two] [huge] [guns] strapped [in] leather holsters [on] his hips.

C Frame Representation (TF-IDF)

```
Prohibiting_or_licensing   : 8.37
Size                       : 6.47
Rebellion                  : 6.31
Local_by_use               : 3.96
Achieving_first            : 3.74
                  ...
```

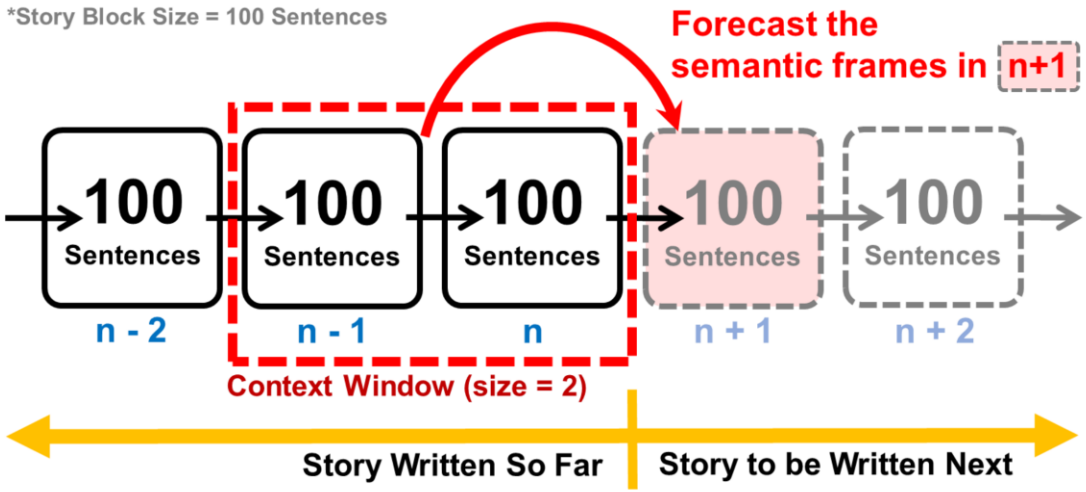
Convert to TF-IDF vector over semantic frames

* IDF is computed on all the corpus!



Semantic Frame Forecast

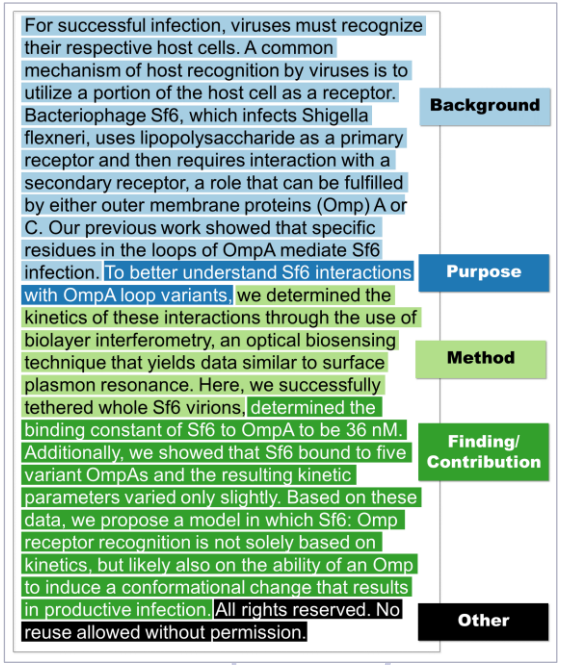
- Using the existing information (block n-1, block n) to predict the **frame representation** in block n+1.
- Cosine similarity for evaluation.



Dataset

- **Bookcorpus** [4]: A total of 4,794 real **fictions** are selected.

- **CODA-19** [5]: A total of 7,962 scholar **abstracts** are selected.

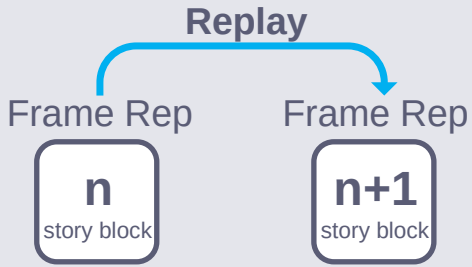


[4] Zhu, Yukun, et al. "Aligning books and movies: Towards story-like visual explanations by watching movies and reading books." *Proceedings of the IEEE international conference on computer vision*. 2015.

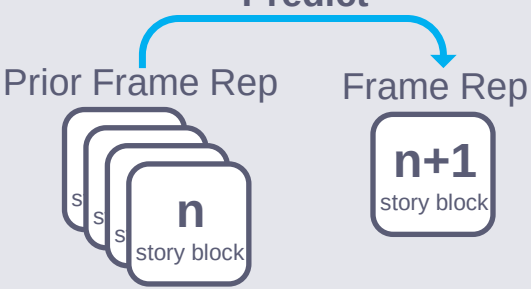
[5] Huang, Ting-Hao, et al. "CODA-19: Using a Non-Expert Crowd to Annotate Research Aspects on 10,000+ Abstracts in the COVID-19 Open Research Dataset." Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020. 2020.

Models

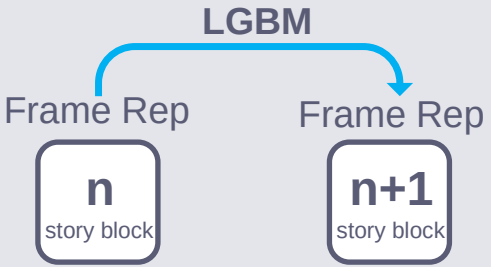
1 Replay Baseline



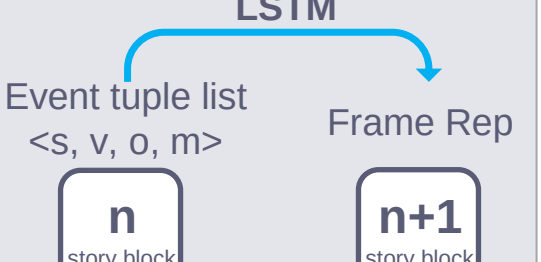
2 Prior Baseline



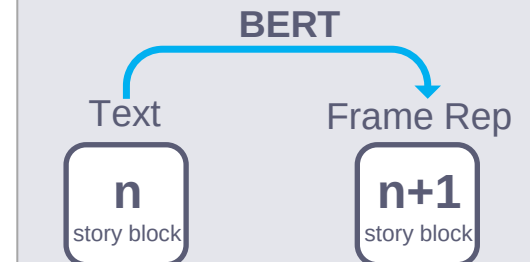
3 LGBM + Frame



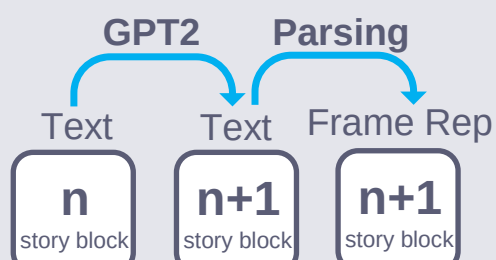
4 Event-Rep [6]



5 BERT



6 GPT2



[6] Martin, Lara, et al. "Event representations for automated story generation with deep neural nets." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 32. No. 1. 2018.

Evaluation

** Cosine Similarity*

Feature	Model	Block Size									
		5	10	20	50	100	150	200	300	500	1000
-	Replay Baseline	.0654	.0915	.1237	.1737	.2163	.2448	.2665	.3000	.3462	.4155
-	Prior Baseline	.2029	.2435	.2857	.3389	.3754	.3962	.4105	.4302	.4528	.4776
Frame	LGBM	.2072	.2506	.2967	.3564	.3995	**.4255**	**.4441**	**.4711**	**.5048**	**.5510**
Event	Event-Rep	.2111	.2541	.2994	.3532	.3929	.4126	.4280	.4453	.4626	.4792
Text	BERT	**.2172**	**.2611**	**.3073**	**.3637**	**.4012**	.4229	.4371	.4559	.4779	.5057
Text	GPT-2	.0519	.0739	.0990	.1402	-	-	-	-	-	-

- **BERT+text** works better when story block is smaller.
- **LGBM+frame** performs better when story block is larger.

Evaluation (cont.)

** Cosine Similarity*

Feature	Model	Block Size									
		5	10	20	50	100	150	200	300	500	1000
-	Replay Baseline	.0654	.0915	.1237	.1737	.2163	.2448	.2665	.3000	.3462	.4155
-	Prior Baseline	.2029	.2435	.2857	.3389	.3754	.3962	.4105	.4302	.4528	.4776
Frame	LGBM	.2072	.2506	.2967	.3564	.3995	**.4255**	**.4441**	**.4711**	**.5048**	**.5510**
Event	Event-Rep	.2111	.2541	.2994	.3532	.3929	.4126	.4280	.4453	.4626	.4792
Text	BERT	**.2172**	**.2611**	**.3073**	**.3637**	**.4012**	.4229	.4371	.4559	.4779	.5057
Text	GPT-2	.0519	.0739	.0990	.1402	-	-	-	-	-	-

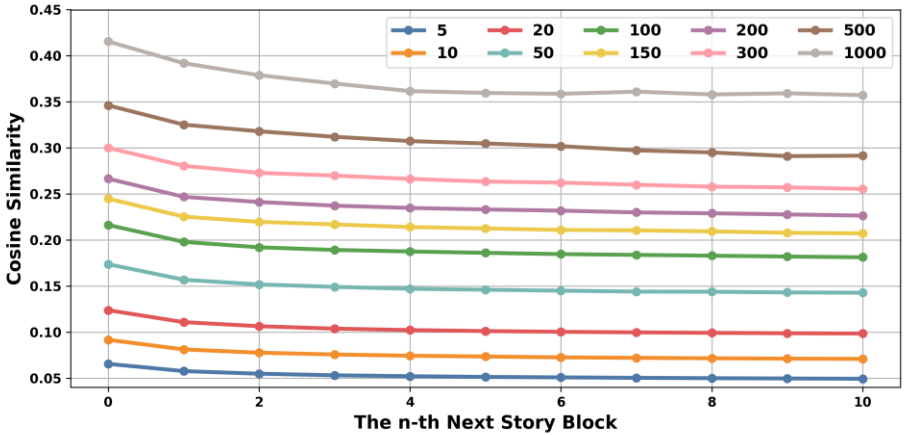
- **Event-Rep** works better in short stories.
- **Prior** is a robust and strong baseline.
- **GPT-2** is not effective => GPT2 fails to maintain coherence among sentences or events.

Is it harder to predict the farther future?

** Cosine Similarity*

Feature	Model	Block Size									
		5	10	20	50	100	150	200	300	500	1000
-	Replay Baseline	.0654	.0915	.1237	.1737	.2163	.2448	.2665	.3000	.3462	.4155

- **Yes! Replay** tells us that things happen now will likely to happen again **shortly**.



Does a longer history help?

- Yes, **LGBM+frame** can **benefit** from using a longer history!

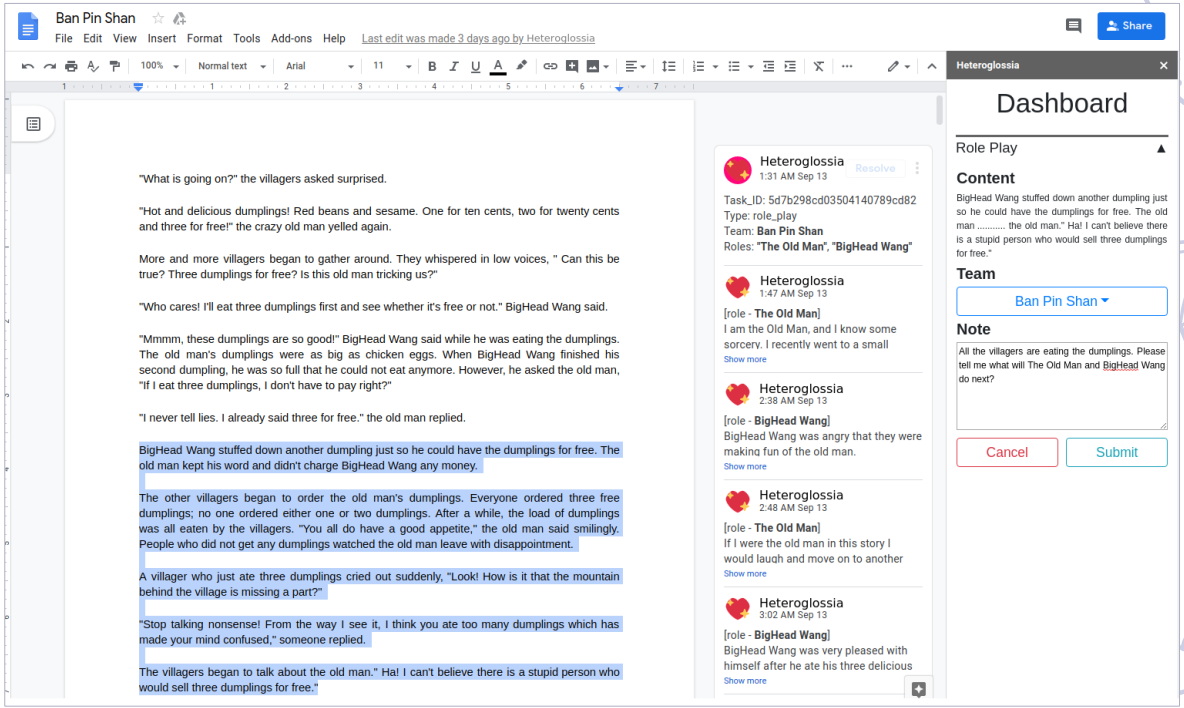
** Cosine Similarity*

window	Feature	Model	Block Size 20	50	100
1	Frame	LGBM	.2967	.3564	.3995
	Text	BERT	**.3073**	**.3637**	**.4012**
2	Frame	LGBM	.2989	.3590	**.4029**
	Text	BERT	**.3081**	**.3625**	.4002
5	Frame	LGBM	.2989	.3617	**.4065**
	Text	BERT	**.3082**	**.3618**	.3985

Increase!!

How do we show the prediction to writers?

- **Heteroglossia** [7] suggests follow-up ideas as a short paragraph using a crowdsourcing method.



[7] Chieh-Yang Huang, Shih-Hong Huang, and Ting-Hao Kenneth Huang. 2020. Heteroglossia: In-Situ Story Ideation with the Crowd. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20).

What to display?

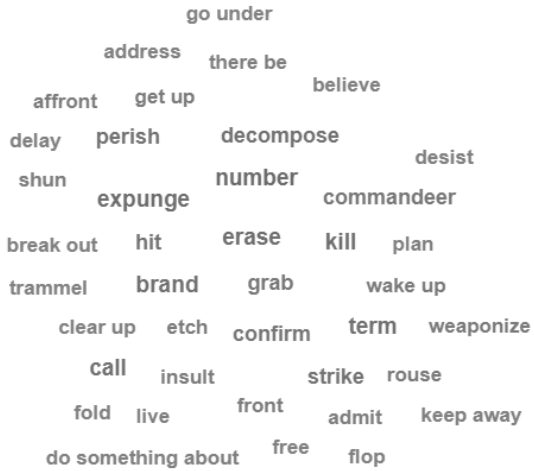
- Semantic concepts!
- Importance!

How to combine these information together?

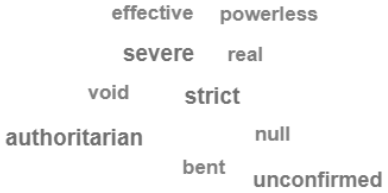
A Possibility: Wordcloud

Noun

Verb



Adjective



Wordcloud Visualization

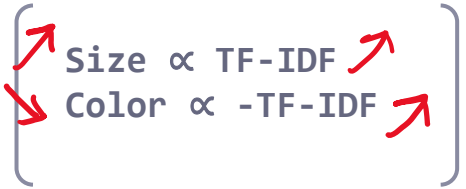
A Top 30 Frames

```
Ineffability              : 16.04
Biological_entity         : 7.31
Be_subset_of              : 4.88
Negative_Conditional      : 4.85
Erasing                   : 4.62
                ...
```

B Corresponding 3 Lexical Units

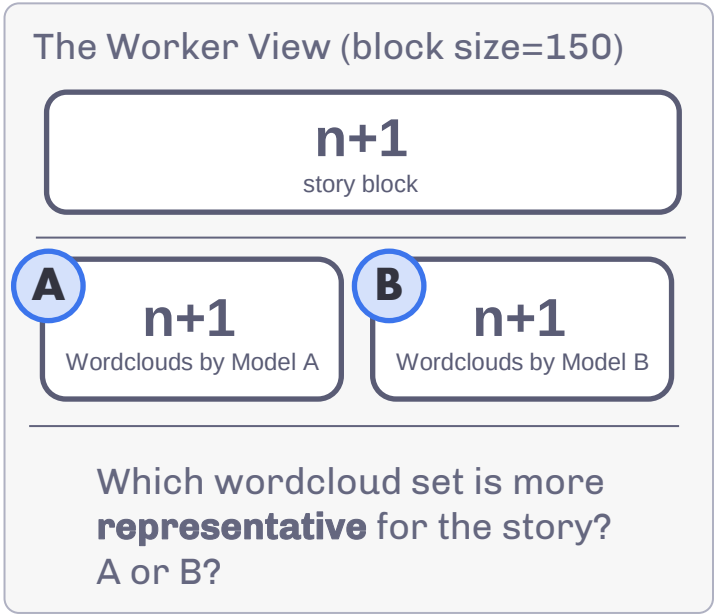
```
magicalness.n / ineffability.n / magic.n
organism.n / life.n / lifeform.n
number.v / subset.n
otherwise.adv
expunge.v / kill.v / erase.v
```

C Compute Size and Color according to the TF-IDF



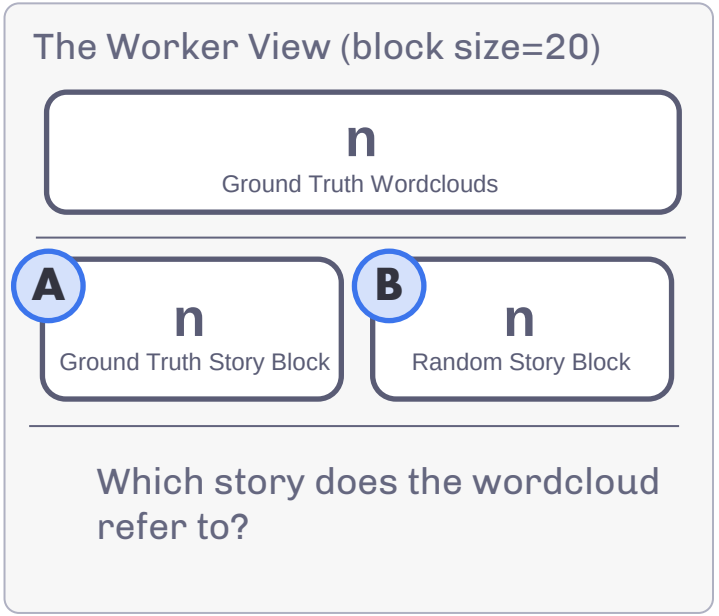
Human Evaluation

- Which model generate a more **representative** word cloud?
- (Ground, LGBM, BERT) = (32, 15, 16)
  ➔ The result makes sense!
- (LGBM, BERT) = (59, 41)
  ➔ LGBM is better!



Human Evaluation

- Does the wordcloud **specific** enough for human to understand?
- YES!! 74% of 50 HITs were answered correctly!



Thanks

Find our code and dataset on GitHub!
https://github.com/appleternity/FrameForecasting

Scan for the code!

Chieh-Yang Huang

Ting-Hao Kenneth Huang