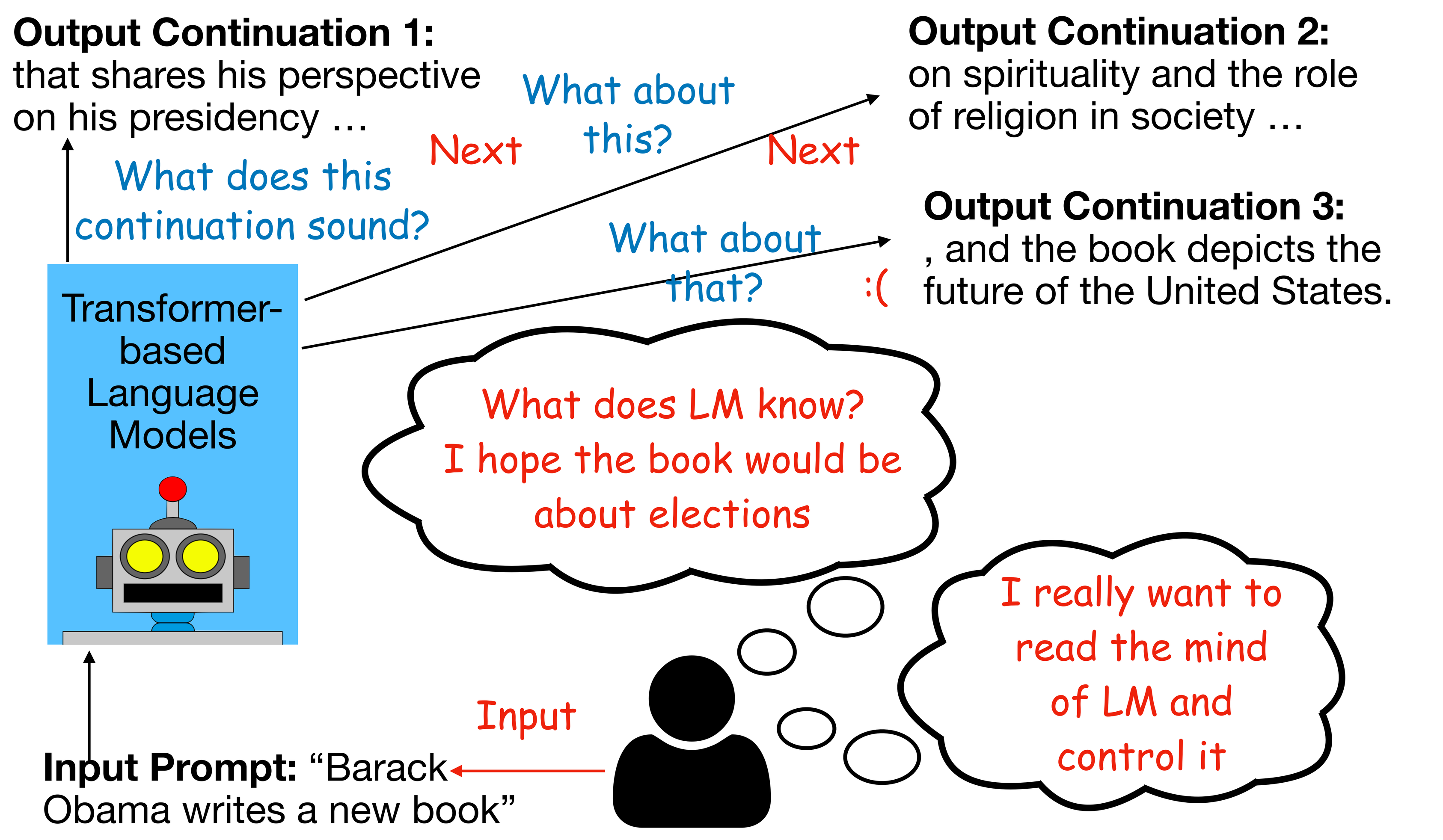
Changing the Mind of Transformers for Topically-Controllable Language Generation

UMASS AMHERST

Haw-Shiuan Chang, Jiaming Yuan, Mohit Iyyer and Andrew McCallum

IESL

Introduction

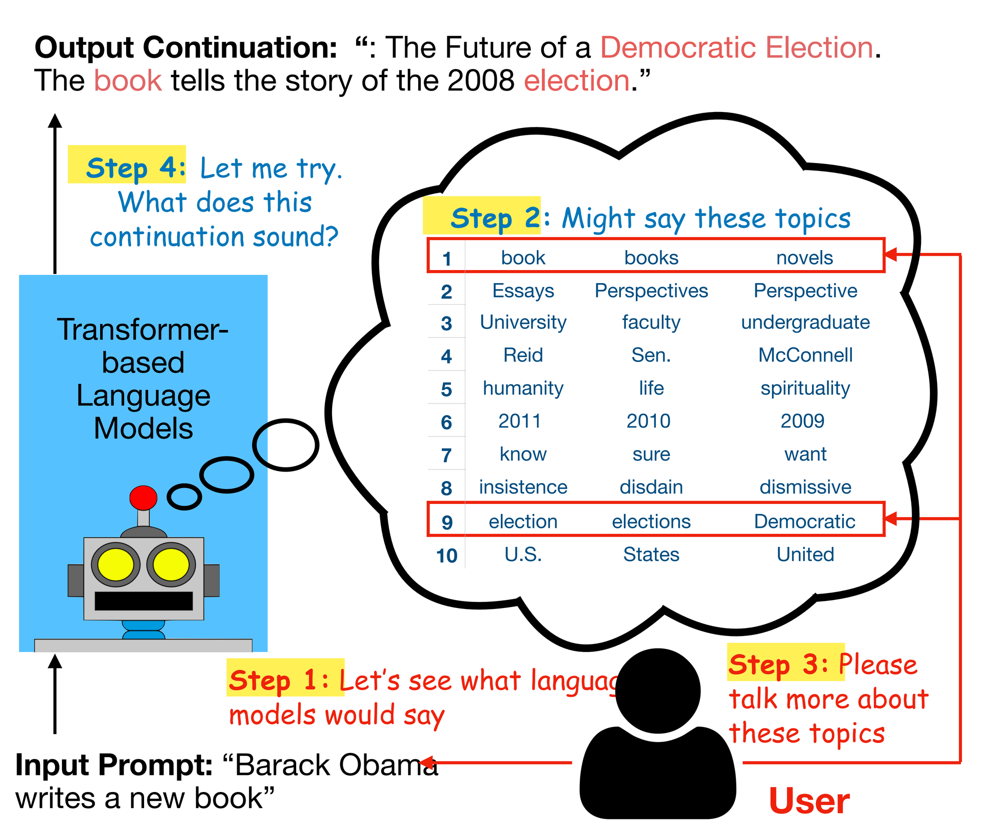


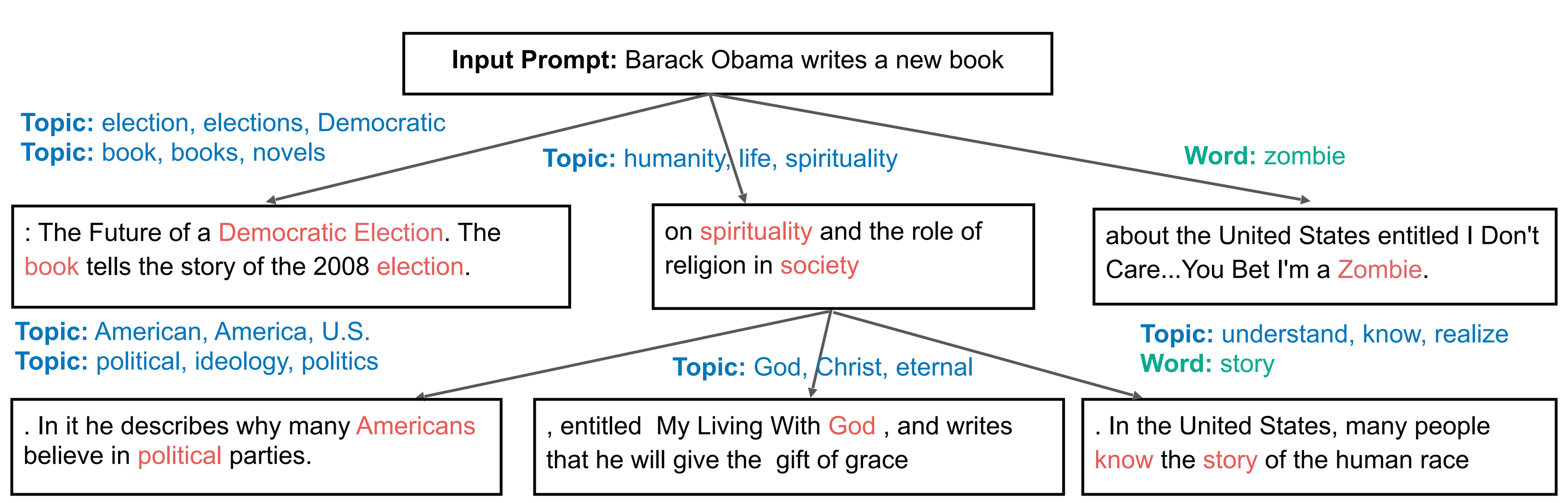
- Our goal is to inject the fine-grained topical preference of users to the language generation model

Our Framework

1. **User:** Specify a prompt
2. **LM:** suggest topics
   - Topics are prompt-dependent and fine-grained
3. **User:** Choose topics
4. **LM:** generate the continuation conditioned on the chosen topics

Repeat -> plot graph

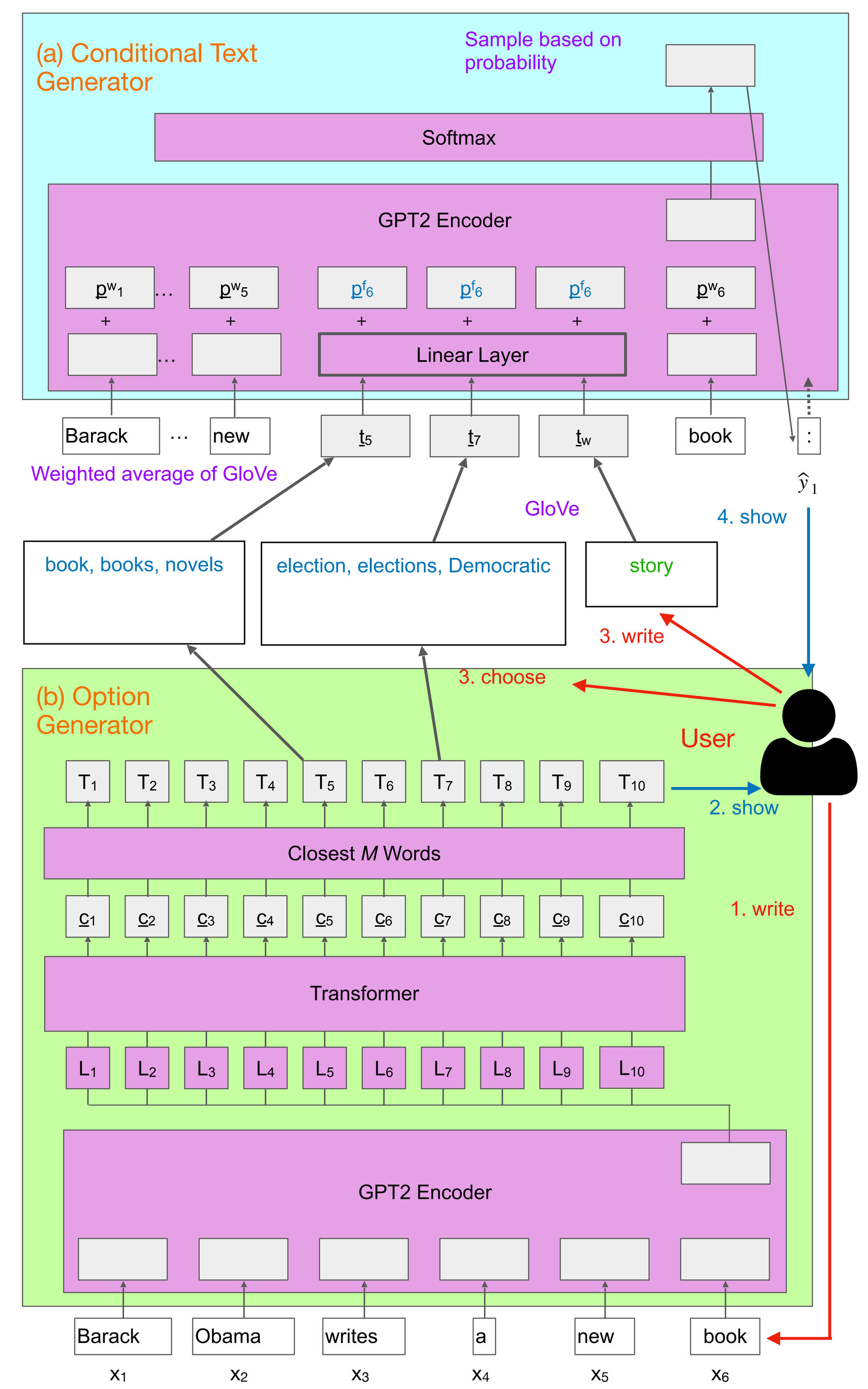




Our Method

Model Architecture (Testing)

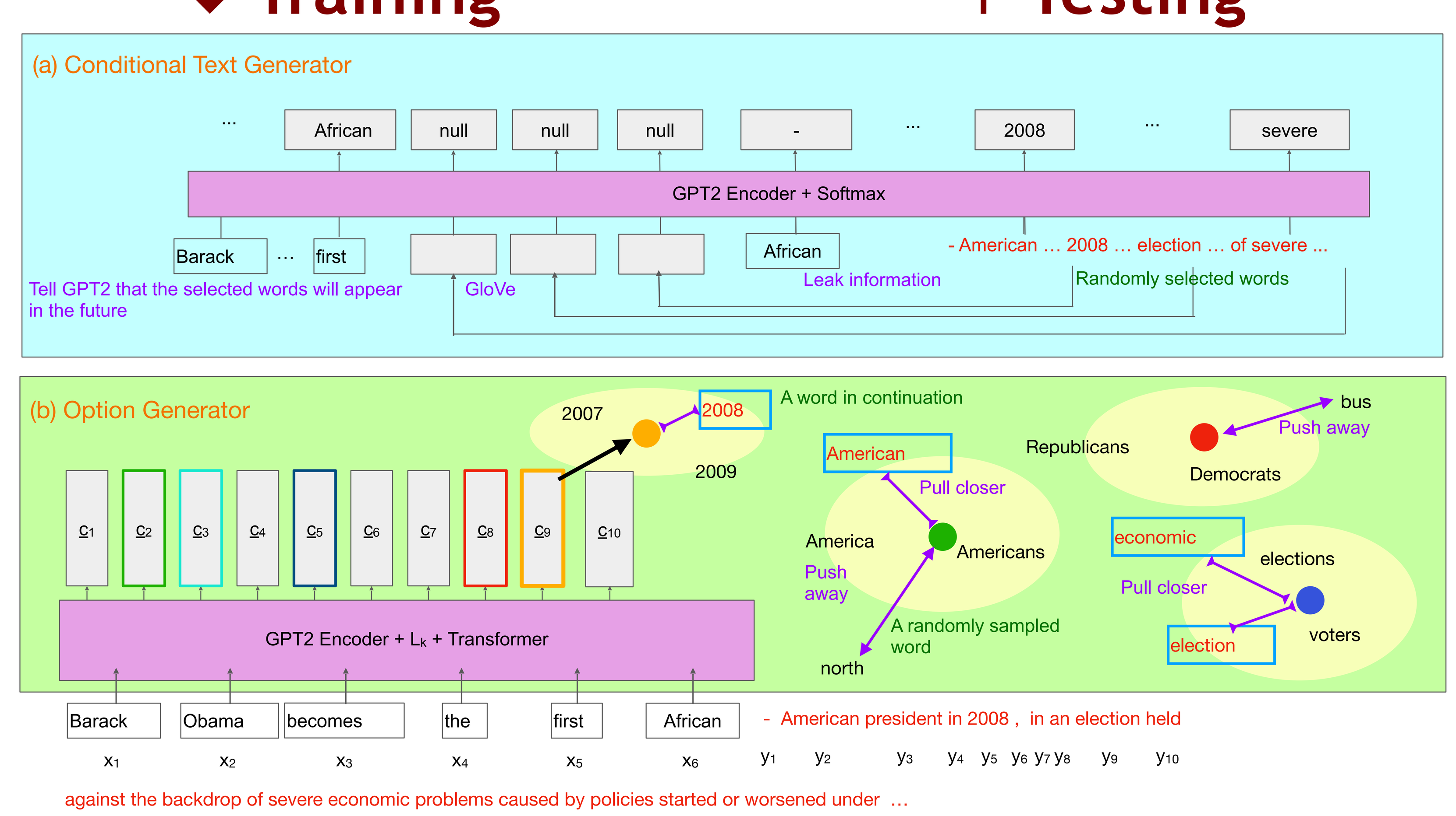
- **Option Generator**
  - Predict the topic embedding in a GloVe space
  - Use 3 words closet to each topic embedding to visualize the topic
- **Conditional Text Generator**
  - Covert the selected GloVe into the size of the hidden state
  - Using top-k sampling



Model Training

- Training without any label data or predefined topics
- **Option Generator**
  - By pulling the topics and the future words closer
    - Topic -> cluster center
  - By pushing the topics and randomly selected words away
- **Conditional Text Generator**
  - Use the GloVe of future words as the condition

↓ Training ↑ Testing



Experiments

Qualitative Comparison

Input Prompt	The study also found that skin cancer nearly tripled in Norway and Sweden since the 1950s.										
LDA-global				Kmeans-local				Ours			
1	population, households	6	company, companies	1	Norway, Sweden	6	also, however	1	research, scientific	6	1980s, 1970s
2	patients, treatment	7	Norwegian, Norway	2	tripled, doubled	7	since, Since	2	tissues, tissue	7	even, though
3	psychology, research	8	story, book	3	nearly, almost	8	Sweden, Finland	3	patients, diagnosis	8	susceptibility, pathogenic
4	police, prison	9	hospital, Hospital	4	cancer, skin	9	study, studies	4	DNA, gene	9	decreased, increased
5	chemical, carbon	10	Icelandic, Iceland	5	1950s, 1940s	10	found, discovered	5	orange, purple	10	Sweden, Norway

Not relevant Redundant

Input Prompt		The study also found that skin cancer nearly tripled in Norway and Sweden since the 1950s.
Generator		Generated Text
Option	Text	
LDA-global	Ours	A study of the Norwegian police has confirmed the cancer case. The law in Norway was the subject of the
Kmeans-local	Ours	The study also found that skin cancer nearly tripled in Norway and Sweden since the 1950s. As well, skin
Ours	PPLM	In this study, a study was conducted conducted in Italy and in Finland. From the 1990s to the 1970s, there
None	GPT2	The study also revealed that only 20% of the deaths in Norway were caused by a sudden cardiac response
Ours	Ours	Recent studies have shown that melanin causes a decrease in genetic susceptibility in people in Norway,

→Not relevant
→Redundant

Option Generator Evaluation

Automatic Evaluation Relevancy Novelty

Scope	Method	Sim	Sim Short	Sim Diff
Global	Sample	14.63	14.42	**0.16**
	LDA	36.86	36.02	-2.82
	Kmeans	**40.65**	**39.91**	-3.40
Local	Sample	41.50	41.23	-12.51
	NNSC	46.70	42.80	-15.94
	Kmeans	47.94	43.89	-16.12
	Ours	**48.38**	**46.29**	**0.45**

Human Evaluation Relevancy Novelty Overall

Scope	Method	L	TP	L&TP
Global	LDA	5.76 ± 0.50	6.24 ± 0.33	5.26 ± 0.31
	Kmeans	6.94 ± 0.36	6.13 ± 0.30	5.96 ± 0.31
Local	Kmeans	**8.65** ± 0.16	5.31 ± 0.50	5.14 ± 0.50
	Ours	7.85 ± 0.25	**6.96** ± 0.26	**6.75** ± 0.28

Human Evaluation Fluency Novelty Overall

Scope	Method	F	NP	A
Global	LDA	3.07 ± 0.17	2.82 ± 0.16	3.06 ± 0.13
	Kmeans	3.65 ± 0.13	3.42 ± 0.14	3.42 ± 0.12
Local	Kmeans	3.71 ± 0.13	3.56 ± 0.15	3.39 ± 0.13
	Ours	**3.85** ± 0.14	**3.64** ± 0.15	**3.67** ± 0.14

Automatic Evaluation Relevancy Novelty Conditionable Diversity

Scope	Method	BLEU	BLEU Diff	Word Hit	Self-BLEU (↓)	Dist-1	Dist-2
Global	Sample	**7.39**	**5.66**	0.34	**9.45**	**47.60**	**86.79**
	LDA	7.19	4.87	**2.01**	13.06	36.02	78.73
	Kmeans	7.12	4.65	1.30	12.23	36.62	81.49
Local	Sample	8.38	2.71	2.93	18.03	35.76	77.00
	NNSC	**8.44**	3.24	2.94	17.20	35.43	76.71
	Kmeans	8.32	3.06	2.96	16.97	35.39	77.10
	Ours	8.38	**5.55**	**3.02**	**15.97**	**36.18**	**78.71**
NA	None	8.50	5.59	-	13.17	39.69	80.17

Conditional Text Generator Evaluation

Relevancy Fluency Diversity Relevancy Fluency

Text Generation Method	Automatic Evaluation						Inference Time	Human Evaluation		
	Relevancy Hit			Quality				Relevancy		Fluency
	Token	Word	Topic	PPL (↓)	Dist-1	Dist-2	s (↓)	Recall	Precision	Score
PPLM	1.48	0.99	0.77	18.49	**40.29**	**80.83**	17.74	30.56 ± 2.96	56.01 ± 4.41	3.83 ± 0.13
Ours	**2.36**	**1.79**	**1.40**	**16.39**	37.98	79.65	**1.02**	**41.46** ± 3.47	**56.41** ± 4.41	**4.07** ± 0.10
GPT2	1.27	0.84	0.64	14.24	39.80	80.22	1.00	24.49 ± 2.77	48.69 ± 4.61	4.15 ± 0.11

Conclusion

- Decompose a novel framework into two novel components
  - Option Generator -> topics are relevant but novel
  - Conditional Text Generator -> Text is fluent and relevant
- Codes are available at https://github.com/iesl/interactive_LM

References

- Haw-Shiuan Chang, Amol Agrawal, and Andrew McCallum. 2021 Extending Multi-Sense Word Embedding to Phrases and Sentences for Unsupervised Semantic Applications. In AAAI 2021
- Paul, Rohan*, Haw-Shiuan Chang*, and Andrew McCallum. Multi-facet Universal Schema. EACL 2021
- Sumanth Dathathri, Andrea Madotto, Janice Lan, Jane Hung, Eric Frank, Piero Molino, Jason Yosinski, and Rosanne Liu. 2020. Plug and play language models: A simple approach to controlled text generation. In ICLR 2020