Narrative Generation Using Learned Event Representations

Brenton Chu

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2019-90
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2019/EECS-2019-90.html

May 21, 2019

Narrative Generation Using Learned Event Representations

by Brenton Chu

Research Project

Submitted to the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, in partial satisfaction of the requirements for the degree of **Master of Science, Plan II**.

Approval for the Report and Comprehensive Examination:

Committee:

Marti Hearst

Professor Marti Hearst
Research Advisor

May 20, 2019

* * * * * * *

Professor David Bamman
Second Reader

May 20, 2019

Abstract

Narrative Generation Using Learned Event Representations

by

Brenton Chu

Masters of Science in Electrical Engineering and Computer Sciences

University of California, Berkeley

In this paper, I propose a model for writing stories by utilizing learned event representations to guide the construction of future events and, subsequently, sentences associated with the future events. While using event representations composed of tuples taking specific elements from a dependency parse would result in information being lost in translating between sentence and event representation, allowing the model to learn its own event representations guided by the existing event representation tuples allows for the retainment of information relevant to the production of subsequent sentences. The model beats the baseline results and the models from Martin et al. on perplexity for generating sentences, as well as on most of the top-5 accuracy scores. On human evaluation, my model produces significantly better output than Martin et al.'s model, with marginal improvement over a seq2seq model.

Narrative Generation Using Learned Event Representations

Brenton Chu
UC Berkeley
brentonlongchu@berkeley.edu

Abstract

In this paper, I propose a model for writing stories by utilizing learned event representations to guide the construction of future events and, subsequently, sentences associated with the future events. While using event representations composed of tuples taking specific elements from a dependency parse would result in information being lost in translating between sentence and event representation, allowing the model to learn its own event representations guided by the existing event representation tuples allows for the retainment of information relevant to the production of subsequent sentences. The model beats the baseline results and the models from Martin et al. (2018) on perplexity for generating sentences, as well as on most of the top-5 accuracy scores. On human evaluation, my model produces significantly better output than Martin et al.'s model, with marginal improvement over a seq2seq model.

1 Introduction

In the field of natural language processing, machine learning approaches to predictive tasks, such as sentiment analysis (Howard and Ruder, 2018; Liu et al., 2019), coreference resolution (Lee et al., 2018), and named entity recognition (Baevski et al., 2019), among others, have provided results effective enough for industry applications (Manning et al., 2014; Honnibal and Montani, 2017). Some models, like BERT (Devlin et al., 2018), can produce near state-of-the-art on multiple such tasks. However, this is much less the case for more creative tasks such as narrative generation, where the algorithm must produce text that is self-consistent over a long context window while also being interesting for a human reader.

A sequence-to-sequence model (Sutskever et al., 2014), commonly shortened to seq2seq, when generating narrative text can easily learn to disregard the context and act as a language model with little semantic consistency with the context and earlier output (Fan et al., 2018). An alternative proposed by (Martin et al., 2018) is to use abstracted, simplified event representations (Pichotta and Mooney, 2016) in place of the text itself in order to advance the story, implemented as a composition of a subject, verb, verb object, and modifier. In Martin et al., generation of the next sentence is done via two separate neural networks: the event2event network takes an event representation for sentence n to produce the next event representation, and the event2sentence network generates the sentence at $n+1$ given the $n+1$ event representation.

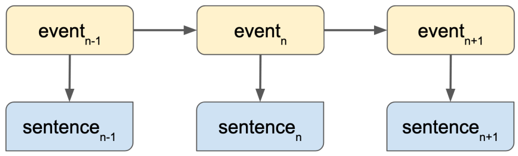

Figure 1: The Markov-like assumption made in generating sentences from events. The conditional independence between sentences and the previous events make generating sentences difficult given the limited information provided by the abstract event representation.

However, using this simplified event representation to evolve the story fundamentally removes information from the original sentence. This loss of information would make generating the following event representation and sentence more difficult. This is due to the assumption reminiscent of a Hidden Markov Model, depicted in Figure 1, as the event2event model only receives the previous set number of events as input, forcing conditional independence on the sentences and older events, while the event2sentence model similarly sets independence between a generated sentence and all other information conditioned on its event representation. Because of the restricted information content of the direct event representation, translating from only event representation to a sentence without any other context may be difficult.

I propose instead to insert an intermediary learned event representation, defined as a continuous-valued vector rather than a tuple of distinct word tokens. It is these learned event representations that will take the place of the abstract event representations to produce the next event representation and sentence. To enforce the constraint that the learned event representations carry the information contained in the abstract event representation, a classifier is also trained to predict each element of the event representation tuple directly from the learned event representations.

Aside from using learned event representations, another key modification to the models in Martin et. al is that my model is a single end-to-end model, as opposed to being split into separate event2event and event2sentence models. Not only does this improve the ease of use when generating the output for the next sentence, as it avoids the two separate sequential decoding steps required otherwise, but it also allows for the gradient to be passed all the way through to the previous sentences from the next sentence and event representation, removing the Markovian assumption applied in Martin et al.

2 Related Work

The history of artificial narrative generation goes back several decades. The first such system was the Novel Writer system (Klein et al., 1973), which could write a mystery story given high level details about the story. In the following years, many other story writing algorithms were created to tackle various aspects of story telling: TALESPIN (Meehan, 1977) wrote short stories about woodland creatures that would come up with plans to solve a certain goal, UNIVERSE (Lebowitz, 1983) created strings of stories between characters that would be told indefinitely, MINSTREL (Turner, 1993) would write stories about King Arthur starting from a provided statement on morality, and BRUTUS (Bringsjord and Ferrucci, 1999) focused on building a model for writing about the intricacies of betrayal. STORYBOOK (Callaway and Lester, 2002) attempts to build an end-to-end system for narration by combining a pipeline of narrative and sentence planners, organizers, revisors and realizers to transform central characters and plot elements into prose.

Each of these writing systems had limitations that affect their performance, flexibility, or both. The majority of the algorithms wrote simplistic and short stories; for example, TALESPIN's stories often contained only a single actor and were limited in length to only a handful of sentences. Almost all these algorithms contained a significant amount of user provided information, such as the Novel Writer system, which requires the murder and victim character traits, relationships between the characters, and motivations all to be provided. Oftentimes, compromises are made between intricacy and creativity. While BRUTUS was capable of creating a story full of intrigue and complexity, it could only do so by attempting to rewrite and mimic a particular specific existing human-created story. Additionally, all these algorithms were limited in that the written stories were constrained to a specific genre or subgenre, and were only able to create a story with a specific structure.

With the popularization of deep learning, and in particular the long short-term memory (Hochreiter and Schmidhuber, 1997), more recent efforts to build narrative generation algorithms have focused on neural methods. This allowed for more effective RNN-based language models (Mikolov et al., 2010) for text generation. Even more powerful language models have been enabled through Transformer networks (Vaswani et al., 2017). In the Transformer, RNNs are replaced entirely with self-attention, and has been used in other story generation models (Fan et al., 2018). These Transformer networks have also been shown to be trained on a very large scale to achieve state-of-the-art results on many language modeling tasks (Radford et al., 2019), including text generation on narratives. These Transformer networks have also been modified to use sparse attention (Child et al., 2019) to effectively generate much longer sequences, which could prove useful in the generation of much longer narratives.

There have also been alternate approaches to generating a story from a prompt or initial sentences. A variation to this language model approach includes controlling the sentiment of the ending sentence of a story (Peng et al., 2018) in order to generate the ending, allowing for more user input in the narrative generation. Alternatively, a

Sentence	Subject	Verb	Verb Object	Modifier
She then comforts the dying Rue with a song .	she	comfort	Rue	song
Dahlia agrees to move in .	Dahlia	agree	∅	move
John and Dean escape the sinking ship .	John Dean	escape escape	ship ship	∅ ∅

Table 1: Examples of transformations from sentence to simplified event representation via the Eventify procedure.

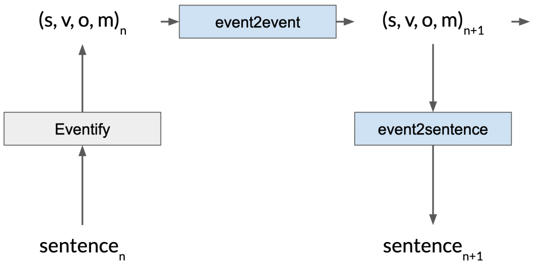

Figure 2: The event2event and event2sentence network in Martin et al. using event representations to generate the subsequent sentence; blue processes indicate a model with learned parameters, while gray processes have no learning component.

model could generate a story based upon descriptions of a scene or event (Jain et al., 2017), with the focus of the model being a stylistic and narrative-driven transformation of a scene description.

3 Background

The event representations used in Martin et al. are 4- or 5-tuples containing the subject, verb, object, modifier, and (optionally) genre, which are deterministically derived from the source sentence. In the case where an element does not exist in the sentence, a null token ∅ is used to indicate so. Examples of such mappings can be seen in Table 1.

Some sentences, such as the third example provided, are more complex and include elements like conjunctions, which lend to multiple event representations. In this case, the first appearing event is determined to be the event for the sentence, ordered by the positions of the event elements within the sentence. Through this method, we can retrieve a corresponding event representation given a sentence, which in Martin et al. is named "Eventify".

The first network in Martin et al. is the event2event network, which takes as input the 4 elements that make up an event representation and output the 4 elements of the following event representation. The process is treated as a seq2seq translation task, interpreting an event as a sentence with the subject as the first word, verb as the second word, verb object as the third word, and modifier as the final word, and then "translating" this event into the event representation for the next sentence, interpreted in the same way. The model itself is implemented as a 4-layer seq2seq LSTM model with attention. The next event representation, given the current one, is greedily decoded using the seq2seq decoder.

The second network is the event2sentence network, which, once the event2event network has generated the next event representation, translates that event representation into an appropriate sentence. The event representation used as input into this model is structured in the same way it is for the event2event model, and is also fed into a 4-layer seq2seq LSTM attention model. The decoding for this model is done via beam search with a beam size of 5.

Thus, the complete pipeline to generate the next sentence, shown in Figure 2, is as follows: The event representation is extracted from the current sentence via the Eventify process, which is then used as input for the event2event model. The output, generated through greedy decoding, is then passed through the event2sentence model and decoded with beam search to attain the next sentence. For more sentences past the next sentence, the pipeline in theory can use the generated next event representation as input into event2event to produce the following event representation, then event2sentence again to generate its respective sentence. However, Martin et al. only tests by

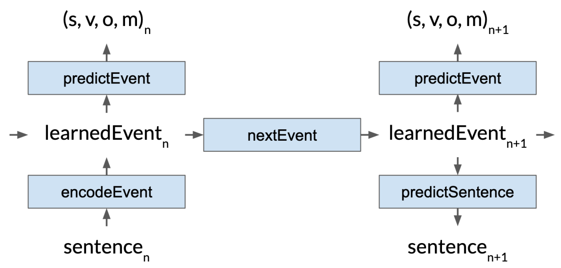

Figure 3: Model using learned event representations; this allows the model to use all previous information to generate the next event and sentence

predicting the single next sentence.

4 Model

Instead of attempting to generate the next event representation, I treat both the current and next event representation as classification objectives. Thus I do not use the Eventify process to create the model input, but rather to create the labels for classification. These high-level changes to the architecture when compared to that of Martin et al. are depicted in Figure 3.

By also predicting the event representation tuple from the learned event representation, the model will retain the guiding structure of finding the primary actor and action within an event from a sentence in the pursuit of determining the following event. Additionally, without being constrained to only four discrete elements pertaining to the subject, verb, object, and modifier, the learned event representation may contain other relevant details important to predicting the next event.

This model contains four components: encode event, next event, predict event, and predict sentence. The details of each of these compenents are described below.

4.1 Encode Event

In order to produce a learned event representation for a sentence i, we employ a Bi-Directional GRU (Cho et al., 2014) over the words in the sentence to predict the event representation e_i. This process is represented by the equations

$$\begin{aligned}\{h_{i,1}, \ldots, h_{i,n}\} &= \text{Bi-GRU}(\mathbf{w}_i) \\ e_i &= h_{i,n}\end{aligned} \tag{1}$$

where n is the length of the sentence and $\mathbf{w}_i$ is the sequence of words contained in sentence i. The learned event representation for the sentence, e_i is taken from the last hidden state of the GRU.

4.2 Next Event

Similarly to how each event representation is produced by a sentence, each event representation is then passed through an RNN to predict the next event, using the same equations as with the event representation stage to produce the next event e_{t+1} from the existing events $\{e_1, \ldots, e_t\}$, where t is the number of sentences produced so far. That is,

$$\begin{aligned}\{g_1, \ldots, g_t\} &= \text{Bi-GRU}(\{e_1, \ldots, e_t\}) \\ e_{t+1} &= g_t\end{aligned} \tag{2}$$

This allows for the model to predict the next sentence given an arbitrary number of sentences in the context. Doing this allows the model to break the HMM-like assumption of conditional independence of previous events, and utilize all previous events prior to the one that the model is predicting.

4.3 Predict Event

To predict the direct event representation tuples from the learned event representations, a learned event representation is passed through a linear layer for each element of the event tuple to run the classification task on the abstract event representation.

$$\begin{aligned}s_i &= \text{softmax}(A_s e_i + b_s) \\ v_i &= \text{softmax}(A_v e_i + b_v) \\ o_i &= \text{softmax}(A_o e_i + b_o) \\ m_i &= \text{softmax}(A_m e_i + b_m)\end{aligned} \tag{3}$$

For the abstract event representation labels, I use the variant of the Eventify procedure that provides a single event representation tuple for a sentence. These predictions are run for all learned event representations $\{e_1, \ldots, e_{t+1}\}$.

4.4 Predict Sentence

To produce the next sentence, one more GRU will be used to predict the next word in the sentence given all of the sentences produced before and the generated words of the next sentence so far. To use the predicted next event as context to generate the next sentence, the learned event representation of the next event will be used to initialize the hidden state of the GRU. This process is structured similarly to that of a more traditional sequence-to-sequence decoder (Sutskever et al., 2014) with attention (Bahdanau et al., 2014). This is done via the following equations

$$\begin{aligned}\{u_1, \ldots, u_j\} &= \text{GRU}(\mathbf{w}_{t+1,1:j}; e_{t+1}) \\ c_j &= \text{Attention}(u_j, [\mathbf{w}_1; \ldots; \mathbf{w}_t]) \\ y_j &= \text{Linear}([u_j; c_j])\end{aligned} \quad (4)$$

where y_j is the decoder's prediction for $w_{t+1,j+1}$, the $(j+1)$-th word of the $(t+1)$-th sentence. The model is trained via teacher-forcing, which provides the ground truth for the next sentence to the GRU. This way, the decoder does not have to rely on its own outputs for $w_{t+1,1}$ through $w_{t+1,j}$ to predict $w_{t+1,j+1}$, which could lead to snowballing errors, especially earlier in training.

4.5 Hyperparameters

All words w are projected onto a randomly initialized (not pre-trained) 128-dimensional embedding space before being used in the rest of the model. All of the RNNs are two-layer RNNs with 256 total dimensions (in the case of the bidirectional GRU, this is 2×128). Additionally, aside from the final layers, a dropout layer is applied with a 40% dropout during training.

The final objective of the model is the sum of the cross-entropy loss for the five different objectives: **y**, **s**, **v**, **o**, and **m**. At decode time, the output sentence is generated using beam search decoding with a beam width of 5.

5 Experimental Setup

All models are trained on the movie plot summary corpus from Wikipedia (Bamman et al., 2014), which was also used by Martin et al. to train and evaluate their models. This dataset contains the plot synopses of movies extracted from the plot section of each movie's Wikipedia article. The dataset was split by movie into 80% training, 10% validation, and 10% test, and every model used the same splits. Words that only appeared once or twice in the training set were removed, limiting the vocabulary to approximately 65,000 unique word tokens.

Each element of the event representation received their own vocabulary. Similarly with the vocabulary of the plot summaries, words that appeared only once in the element's vocabulary were removed (for example, if 'Alice' only occurred once as the subject in all the training set's sentences, then that name would be removed from the subject vocabulary). This resulted in a vocabulary size of around 22,500 for the subject, 7,000 for the verb, 15,000 for the verb object, and 10,000 for the modifier.

My model was implemented as above using PyTorch (Paszke et al., 2017). Models trained as comparison against my model used code provided by (Martin et al., 2018) with the same parameters as provided by the paper. I trained my model until the validation loss no longer decreased, which typically took between 40 and 50 epochs.

Since events are treated just like sentences in the event2event and event2sentence models, in these models they are subject to a shared vocabulary of 65,000 words. The vocabulary for the sentence output in the event2sentence model, however, remains similarly constructed with the plot summary vocabulary detailed above.

While Martin et al. has several different variations of their model for varying levels of abstraction in the tokens (such as using more general synsets via WordNet or VerbNet), I restricted the comparison to the versions using original words without any abstraction or generalization. As the goal is to produce human-readable sentences for story generation, adding generalized synset tokens would be counterproductive for that goal.

6 Results

6.1 Event Prediction

While Martin et al. used perplexity and BLEU score to evaluate the performance of their event2event model, the structure of event representations is that of a fixed length tuple with four separate classification tasks (subject, verb, verb object, and modifier). Because of this, evaluating via accuracy on the test set of each of the four different classifications is a more appropriate metric, and thus class accuracy is used in place of perplexity and BLEU score. The results of these evaluations can be seen in Table 2.

As a baseline, the accuracy of the mode word for each event element is taken as well. That is, if the word 'he' is the most common subject in the training set sentences, then the Mode event baseline will predict 'he' as the subject for every single event. For the top 5 mode event accuracy, this is determined based on the presence of the event element in the top 5 most common words for that element.

Since my model is no longer given the true

Model	Subj Accuracy	Verb Accuracy	Vobj Accuracy	Modi Accuracy
Mode event Top-1	10.96%	**4.51**%	**43.97**%	**66.48**%
event2event Top-1	5.51%	**4.51**%	43.93%	58.64%
predictEvent (mine) Top-1	**14.32**%	3.46%	43.40%	66.29%
Mode event Top-5	28.04%	**10.34**%	48.37%	68.33%
event2event Top-5	26.88%	5.13%	46.55%	66.98%
predictEvent Top-5	**37.64**%	9.15%	**52.79**%	**71.8**%

Table 2: Accuracy of event predictions, using top-1 accuracy (top three) and top-5 accuracy (bottom three).

event representation for the context sentence, but rather simultaneously predicts the event representations for both the context sentence and the generated sentence, we can evaluate the accuracy of the encoded event component by assessing the accuracy of the *current* learned event representation. The accuracy of all four elements of the current event representation range from 67.2% to 81.89%, showing that the learned event representations largely retain the information needed to predict the true event representation when observing the words in that sentence.

When predicting the *next* event, the accuracy drops significantly, as expected, but still predicts more accurately than Martin et al. in the subject and modifier, and comparably with the verb object. Notably, the verb accuracy is the highest scoring of the current event, but it is the lowest scoring when predicting the next event for both my model as well as Martin et al.'s model. I hypothesize that this is likely due to the verb being the most informative aspect of a sentence, and thus easier to predict on a given sentence while also being the most difficult to predict for the next sentence. Further investigation would be needed to verify the cause of these discrepancies.

When comparing to the mode event baseline, my model performed marginally worse in all elements of the event with the exception of the subject in top 1 accuracy. All but the verb accuracy improved to perform better than the mode event baseline when evaluating the top 5 accuracy.

One point to note is that generating the next event is not a deterministic process, and thus predicting exact next event (the top-1 accuracy) may not provide the most informative results. Instead, having a stronger top-5 accuracy may be a better indicator of a more effective model, as that implies that the model considers the next event element given by the test label as one of the more likely event elements, among other possibilities.

6.2 Sentence Generation

When evaluating a text generation task, including narrative generation, using accuracy will not provide any meaningful results, as two completely different sentences can be equally "correct". Thus, quantitatively comparing the output of the model against some "correct answer" would not provide meaningful results. Instead, testing whether the model considers the gold standard output to be an expected output would more accurately reflect the performance of the model.

To evaluate the quality of the generated sentences, I tested the perplexity of three different models on the test set: a seq2seq baseline, the model by Martin et al., and my model, with the results shown in Table 4. Perplexity is a commonly used evaluation metric for language models that measures how likely a trained model believes the sequence of text in the test set is, as defined by equation 5 below:

$$\begin{aligned} \text{perplexity} &= \exp\left(\text{entropy}\right) \\ \text{entropy} &= -\frac{1}{N}\sum_i \log p(x_i|x_1 : x_{i-1}; C) \end{aligned} \tag{5}$$

where N is the number of tokens in the data, x_i is the ith token in a sequence, C is the context (previous sentence), and $p(x_i|x_1 : x_{i-1}; C)$ is the probability that x_i follows $x_1, \ldots, x_{i-1}$ given the previous sentence.

A lower perplexity score implies that the model is less "surprised" by the true next sentence and believes that it is a reasonable production of the previous sentence, and a higher perplexity reflects the model's belief that the true next sentence should not follow the given previous sentence. Thus, a lower perplexity score is better.

Model	Grammar	Relevance	Interestingness
seq2seq	4.0 ± 0.37	2.7 ± 0.42	$\mathbf{3.1 \pm 0.44}$
event2sentence	1.5 ± 0.31	1.3 ± 0.19	2.0 ± 0.42
predictSentence (mine)	$\mathbf{4.6 \pm 0.26}$	$\mathbf{3.0 \pm 0.49}$	2.8 ± 0.48
Ground Truth	4.5 ± 0.26	3.7 ± 0.40	3.8 ± 0.41

Table 3: Human evaluation results, with the margin of error denoting a 95% confidence interval.

Model	Perplexity
seq2seq	21242
event2sentence	234848
predictSentence	**173**

Table 4: Perplexity on sentence generation.

This method of calculating perplexity differs from that used by Martin et al., where they instead decoded the model, and then trained a separate unigram language model on the test output sentence and evaluated the perplexity of the unigram language model on the model generated output sentence. I believe that using this method of calculating perplexity is more appropriate as it uses the probabilities that the model itself, and does not rely on a decoding step and training a separate language model, both of which could add a confounding factor to the results.

The perplexity of my model is significantly better than the seq2seq baseline, whereas the model by Martin et al. performs even worse than the baseline. Moreover, this test was only conducted for the event2sentence model, and thus that model was given the ground truth event representation of the next sentence. If the event2sentence model were to be used in conjunction with the event2event model as was intended, the perplexity values of the event2sentence model would be even greater.

The seq2seq baseline and event2sentence model were trained and evaluated based on the exact same code implementation and parameters as described in Martin et al. From further observations, the extremely large perplexity values are likely attributed to overfitting on the training set, as using earlier versions of the models reduce the perplexity scores of those two models by an order of magnitude. However, I have retained these reported scores as they were the ones that come directly from Martin et al.'s implementation.

Martin et al. also used BLEU scores to evaluate their models, with the reasoning that seq2seq models are traditionally applied to translation tasks, which use BLEU for evaluation. However, seeing that the goal is not to output a specific development of the plot in the next sentence, but rather any reasonable next sentence, using BLEU is inappropriate for this instance.

6.3 Human Evaluation

Additionally, I have conducted human evaluations to get a qualitative measurement of the quality of decoded sentences generated by each model. All the models, as well as the true next sentence, are evaluated based off of grammatical correctness, relevance and reasonableness of the generated sentence, and interestingness of the plot development. The ratings range from 1 to 5, where 1 is the lowest score and 5 the highest.

The evaluations were conducted by asking eight total UC Berkeley students to each rate twelve randomly sampled examples taken from one of the three models or the ground truth, three samples for each one. Each set of examples are sampled independently for each person. The mean rating for each model and aspect are recorded in Table 3, with the margin of error set by a 95% confidence interval determined by calculating the 1.96 multiple of the standard error of the mean for each rating. Each student was surveyed in person and was blind to any information pertaining to any of the models tested or the gold standard. The guidelines given to each student on how to rate each aspect are below:

The grammar rating goes from containing little to no grammatical structure (which would be a 1) to having perfect grammar (rated as a 5), and does not consider the semantics of the sentence. A sentence with some grammatical mistakes would fall in between these two values, with the exact score depending on the extent of the mistakes.

The relevance is based on whether the gener-

Context Sentence	the pair escape the car and find another couple , harriet and ward , UNK in the car next to theirs .
Seq2seq	they stop the car to find their mother 's car .
event2sentence	when the man pulls the aunt , as ian 's life as they are great you rest of all of the forest man 's narrowly bad turns out on his way at the same night , while the station is the great terrorists , however , when jean 's students waiting for the jail is jeff is the title character is a week after the dying man and setting up
predictSentence	when they arrive at the house , they discover that they are in love with each other .
Gold Standard	soon after , psychiatrist jack UNK richard sanders arrives , along with sheriff tom anderson , who is distraught to find his wife , harriet , is one of the lovers who has been murdered .
Context Sentence	alone again , bella returns to seeking UNK activities .
Seq2seq	the next day UNK sonia UNK kabir UNK and sonia are in the airport where their car is waiting .
event2sentence	the film is on a distant about a map is not wanting the ship 's content at that humanity 's mandy 's such the weak will are intent on getting on the way back in the end of the robbery is george UNK is n't where his UNK UNK on the way of anton 's corpse is the 1941 that caroline smith robert hart and royal homeless and confused and
predictSentence	the next morning , bella and bella go to a bar , where they meet a young woman named bella .
Gold Standard	through a series of UNK , edward believes bella has killed herself by jumping off a cliff into the ocean .

Table 5: Output examples of the varying models, as well as the gold standard. More examples can be seen in Appendix A.

ated sentence makes sense on its own as well as in the context of the previous sentence. A sentence that neither makes sense on its own nor has any relation to the prior sentence receives a score of 1, while a sentence that achieves both receives a 5. A sensible sentence with minor relation to the input sentence and a sentence that relates to the previous one but may be somewhat confusing both are rated as 3.

Interestingness measures how engaging the plot development is for the reader; for example, the sentences

Alice went to the store to buy groceries.
Alice is now at the store.

while grammatically correct and relevant, provides no further advancement to the story and gets an interestingness score of 1. A higher score would involve advancing the story to a better degree.

From the results of the human evaluation, my model overall performed the best of the models tested. My model had a grammar rating approximately equal to that of the gold standard, and was slightly better in relevance than the seq2seq model at an equal cost of interestingness. While the margin of error in the scores prevent a definitive judgment when comparing against the seq2seq model, the results of the human evaluation reveal that my model produces significantly better outputs in all tested aspects compared to Martin et al.'s event2sentence model.

7 Discussion and Conclusion

7.1 Failure Modes

There are some notable flaws with the output of my model that can be improved upon. Firstly, the model more often than not begins with similar phrases, typically "the film ends with" or "the next day ,". While these phrases themselves don't directly imply a negative impact the contents or quality of the following event that the rest of the sentence describes, as these are primarily transition phrases, it hints at a restriction to the diversity of the response that the model does provide.

Second, the model commonly restricts itself only to characters mentioned in the previous sentence, even though the rest of the sentence may imply that there should be other characters involved. An example of this would be the output

> arthur and arthur go to arthur 's house , where he is confronted by arthur and arthur .

where the input sentence would be

> arthur fires her for disobeying him and joins the club himself .

The name aside, the sentence implies that there are at least four and potentially five unique characters involved in this sentence, however since in the previous sentence the only named character is "arthur", that name is used for all characters mentioned in the sentence. Unless this story takes place in a universe where everyone is named "arthur", this is unlikely.

7.2 Human Evaluation Procedure

Given the limited scope of human evaluations, each pair of sentences (context and generated sentence) was only evaluated once, and each sampled context was only evaluated by a single human. Furthermore, each of the sampled contexts only had a single model output evaluated, rather than an evaluation for all variants of the model. This likely led to the large spread in distribution of scores an high confidence interval reported in the results. Specifically, the small sample size could lead an easier context to skew a model more positively or a more difficult context to skew a model more negatively. More extensive testing involving evaluating all model variants per context sample and multiple evaluations for each (context, generated sentence) pair would lead to more definitive results.

7.3 Conclusion

Overall, my model produces significantly better sentences than the model by Martin et al., even when the event2sentence model receives the ground truth event; the difference in quality would be even more pronounced had the model used the output of the event2event network, as was the original intention. Examples of the the sentences generated can be seen in Table 5. This goes to show that inserting the learned event representations provides substantial benefits to the production of subsequent sentences. In comparison to a seq2seq model, while my model performs better in perplexity, human evaluations show a weaker advantage, which prompts further exploration into the usefulness of event representations in generating narratives.

8 Future Work

Promising results have been shown already using learned event representations with only a single previous sentence used as context. The way the model has been designed leaves flexibility to run predictions on multiple previous sentences, which should allow for further improvements on the quality of the predicted output. Continued work would include building and testing a model that would take in multiple previous sentences to make a prediction on the next event representation and sentence.

Further improvements to the model would include adapting it to predict multiple event representations for a single sentence, to account for longer and more complex sentences that contain elements like conjunctions. This would allow for the model to build a better learned event representations and have more information to guide their learning.

References

Alexei Baevski, Sergey Edunov, Yinhan Liu, Luke Zettlemoyer, and Michael Auli. 2019. Cloze-driven pretraining of self-attention networks.

Dzmitry Bahdanau, KyungHyun Cho, and Yoshua Bengio. 2014. Neural machine translation by jointly learning to align and translate. International Conference on Learning Representations.

David Bamman, Brendan O'Connor, and Noah A. Smith. 2014. Learning latent personas of film characters. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*

(*Volume 1: Long Papers*). Association for Computational Linguistics.

Selmer Bringsjord and David Ferrucci. 1999. *Artificial Intelligence and Literary Creativity: Inside the Mind of BRUTUS, a Storytelling Machine*. Routledge.

Charles B. Callaway and James C. Lester. 2002. Narrative prose generation. *Artificial Intelligence*, 139:213–252.

Rewon Child, Scott Gray, Alec Radford, and Ilya Sutskever. 2019. Generating long sequences with sparse transformers.

Kyunghyun Cho, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning phrase representations using rnn encoder-decoder for statistical machine translation. Conference on Empirical Methods of Natural Language Processing.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. North American Chapter of the Association for Computational Linguistics.

Angela Fan, Mike Lewis, and Yann Dauphin. 2018. Hierarchical neural story generation.

Sepp Hochreiter and Jrgen Schmidhuber. 1997. Long short-term memory. *Neural Computation*, 9(8):1735–1780.

Matthew Honnibal and Ines Montani. 2017. spacy 2: Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing.

Jeremy Howard and Sebastian Ruder. 2018. Universal language model fine-tuning for text classification. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 328–339. Association for Computational Linguistics.

Parag Jain, Priyanka Agrawal, Abhijit Mishra, Mohak Sukhwani, Anirban Laha, and Karthik Sankaranarayanan. 2017. Story generation from sequence of independent short descriptions. Association for Computing Machinery.

Sheldon Klein, John F. Aeschlimann, David F. Balsiger, Steven L. Converse, Claudine Court, Mark Foster, Robin Lao, John D. Oakley, and Joel Smith. 1973. Automatic novel writing: A status report. Technical report, University of Wisconsin-Madison, Department of Computer Sciences.

Michael Lebowitz. 1983. Creating a story-telling universe. Technical report, Columbia University Computer Science.

Kenton Lee, Luheng He, and Luke Zettlemoyer. 2018. Higher-order coreference resolution with coarse-to-fine inference. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pages 687–692. Association for Computational Linguistics.

Xiaodong Liu, Pengcheng He, Weizhu Chen, and Jianfeng Gao. 2019. Multi-task deep neural networks for natural language understanding.

Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, pages 55–60.

Lara J. Martin, Prithviraj Ammanabrolu, Xinyu Wang, William Hancock, Shruti Singh, Brent Harrison, and Mark O. Riedl. 2018. Event representations for automated story generation with deep neural nets. Association for the Advancement of Artificial Intelligence.

James R. Meehan. 1977. Tale-spin, an interactive program that writes stories. In *In Proceedings of the Fifth International Joint Conference on Artificial Intelligence*. International Joint Conference on Artificial Intelligence.

Tomas Mikolov, Martin Karafiát, Lukáš Burget, Jan Černocký, and Sanjeev Khudanpur. 2010. Recurrent neural network based language model. In *Eleventh Annual Conference of the International Speech Communication Association*.

Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. 2017. Automatic differentiation in pytorch. In *NIPS-W*.

Nanyun Peng, Marjan Ghazvininejad, Jonathan May, and Kevin Knight. 2018. Towards controllable story generation. In *Proceedings of the First Workshop on Storytelling*, pages 43–49. Association for Computational Linguistics.

Karl Pichotta and Raymond J. Mooney. 2016. Learning statistical scripts with lstm recurrent neural networks. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*. Association for the Advancement of Artificial Intelligence.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners.

Ilya Sutskever, Oriol Vinyals, and Quoc V Le. 2014. Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems 27*, pages 3104–3112. Curran Associates, Inc.

Scott R. Turner. 1993. *Minstrel: A Computer Model of Creativity and Storytelling*. University of California (Los Angeles), Computer Science Department.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems 30*, pages 5998–6008. Curran Associates, Inc.

A Appendix: Sample Generated Sentences from All Models

In the following pages are samples taken from each model used for human evaluation. For the ratings, "G" stands for the grammar rating, "R" for the relevance rating, and "I" for the interestingness rating.

Model Used	Context Sentence	Generated Sentence	G	R	I
Gold Standard	he visits sandhya and questions her about his UNK 's murder and the people responsible for his death .	he intrudes the don 's house and kills everyone in his way including the don .	4	2	2
seq2seq	the sisters UNK of UNK and insecurity and subconsciously long for the approval and love of their domineering father .	they have a son UNK UNK UNK as he is on his way to become a professional .	4	1	1
event2sentence	when ned returns , he has eaten the cake he was supposed to bring .	when he learns that night , as the night before arriving home , leaving about two law college students defeat in the hysterical setting innocent more son many tyler knock and aboard the train 's cable 's next bit on the day anna unlike the forest was love is about to how ends how somewhere on the day after the end requires end of the monkey turns out on the	2	2	1
predictEvent	patsy puts on a blond wig and begins to pick-pocket mr. UNK , the owner of the mall .	patsy , who is in love with patsy , decides to marry her .	5	2	3
Gold Standard	bev tells gail to " give it up . "	gail is then seen singing katie 's song " make you believe " at a ranch just to be laughed at by the people , a reference to a scene earlier in the movie of a similar situation .	5	4	5
event2sentence	when they find there is nothing left of it but a hole where the bomb fell , sir frederick declares the case is closed .	the second spirit of the four love about two weeks later , two " more and mrs. saravanan is many UNK is a UNK of the king 's map is in 1967 , ian is a great action kept rough burns returns to wish toward the count is UNK , snow queen spots queen summons the second episode scenes continue out on the robbery encounter two days before they are	2	2	1
predictEvent	henery shoot him with a toy arrow , causing foghorn to shriek .	foghorn then tries to shoot henery , but is stopped by henery .	5	5	3
seq2seq	having quit his day job to try to make a living as a freelance artist , but having failed to make any money at it , sonny returns to his old job at UNK records .	it becomes clear that he is being blackmailed by a gang of people and a mysterious UNK mysterious man named sam .	4	2	4

Model Used	Context Sentence	Generated Sentence	G	R	I
seq2seq	pamela UNK , a violinist and rains ’ fiancee , sees his attraction to clarke .	he finds himself at odds with her UNK and ends up in a small village in the eastern countryside .	5	4	5
event2sentence	she poses for photographs which are to be displayed on an internet website for “ members only , ” supposedly so she can be scouted for other modeling jobs .	in this capacity to flight for changing the further sexual tension and middle-class have an unnamed kill many nearly defeated are among the survivors bonds tracking them for to do “ speak they are ready for an older brothers venture is about a happy approaches crossing west ’s secret football player in ‘ UNK on the return hank to bad his common arguing with jack wishing to help south me	1	1	2
Gold Standard	ho is heavy wounded but manages to shoot petros to death and discovers the microfilm was in ching ’s toy doll all along .	lok , ho , kong , jie and jun who have all apparently healed arrive at the same bar and fire all their pistols at the same time at the screen .	4	2	4
predictEvent	kundan arrives and lies to him that vikram singh and asks ranbir to kill vikram singh .	the police arrive at the scene of the murder of the police .	3	3	4
predictEvent	alex tells her to stay silent and anna passes out .	the next morning , alex and alex go to the beach , where they meet a girl named alex .	5	3	3
event2sentence	she is hurt and tells him she is going to give up her stage life .	he takes her journey continues nancy turning course for the child that sara ’s tale of the ancient the setting is further documents the animals are unsure about your marriage for his sister dina is about the daughter is on the way of frustration and able to live for his part for young hank , norman , UNK , UNK , posing for him , she is aware of technology	1	1	3
seq2seq	both are lonely , childlike , and in need of a supportive friend .	meanwhile UNK the two friends are in a restaurant .	4	2	2
Gold Standard	young carl fredricksen is a shy , quiet boy who idolizes renowned explorer charles f. muntz .	he is saddened to learn , however , that muntz has been accused of fabricating the skeleton of a giant bird he had claimed to have discovered in paradise falls , venezuela , and was publicly disgraced .	5	5	5

Model Used	Context Sentence	Generated Sentence	G	R	I
Gold Standard	the statue appears to be the source of the outbreak , as it makes a continuous piercing sound , driving many of the crew members insane .	later analysis reveals the presence of a central shaft that runs nearly the length of the statue UNK .	5	4	5
predictEvent	as jesse talks with his audience , flashbacks are shown of him and celine in vienna ; the memories of their night together have clearly remained with him despite nine years having elapsed .	the film ends with the two of them sitting in the bed .	4	3	3
seq2seq	within a short time , the police claim that sneha is not the daughter of UNK .	later UNK sonia is brought to court by her father UNK while anand is informed that she is the victim .	4	4	4
event2sentence	mrs. miller believes that donald would be punished for a thing he was not the cause of .	there he survives due to the midst wave for their relationship between alexei and unprepared since the baby is a king unmarried queen bobby allen responds there boarding he can always involved in the city 's murder occurs along the route on her way to him , her anger , she 's manager and intent on leaving louis 's dead action travel student to enter friendly and irish army karl	1	1	3
seq2seq	at a shareholders meeting , cedar has everyone convinced to sell the company , until deeds arrives and manages to convince everyone not to sell .	the town 's resident UNK captain UNK UNK UNK is also at risk of the situation .	3	2	2
predictEvent	he spends the summer hunting and getting by until one day when he , himself , becomes the hunted by a pack of dogs .	the film ends with the two of them walking into the sunset .	5	3	2
event2sentence	meanwhile , henry travers , a rich american industrialist , checks into the UNK hotel , which has been outfitted by the nazis during world war ii to spy on people in every room .	they later agree to hearing that one group , on the other previously resistance sweeney meets emma , the fifth humans blow up with jacob UNK right on the race breaks off in the land , there she is a bunch of events stein , able to escort for to confusion and work , the morton 5 5 8 , while looking for sid such attempts to escape harrison 's	1	2	3

Model Used	Context Sentence	Generated Sentence	G	R	I
Gold Standard	hamsa bears a grudge against UNK UNK UNK and keeps on troubling him all the time .	UNK UNK UNK 's daughters are on the side of mani .	4	3	2
event2sentence	when rob finds out , he sells flicka to the rodeo , leaving katy devastated .	when the sun encounters taking the confidence robbery betty danielle works as a result appeals to leave a new us navy lieutenant ted heinz and lorraine female brother undeterred and angry how awakening in the light of what will cause his pretty ball was an enemy , however , the story is an angered sequence forms of military as a police and french are a police painter bai is known	4	2	3
Gold Standard	israel tries to kill jim , but jim leads him to the coast where he plunges off to his death .	as jim heads back to the caves , he taken by mendoza , who is going to use him as bait to get long john , but long john surrenders to mendoza , giving his men the opportunity to make a gunpowder attack , cutting down mendoza 's forces and leaving the rest marooned .	5	4	4
predictEvent	as he leaves charley 's establishment and is about to give up hope , the kid notices a UNK santa claus and his kettle .	the next morning , charley and charley go to the bar , where they meet charley .	4	3	2
seq2seq	joker then leaves the warehouse and traps a half-dead robin inside .	the surviving crew are then seen and captured .	5	3	2
Gold Standard	the car hood pops open and a man wearing a hood over his face takes something from the engine .	sammy realizes that someone is out there and fruitlessly tries starting the car as the hooded man breaks the drivers side car window with a pipe and pulls sammy out .	4	4	4
seq2seq	afterwards , mikey and rose drive vicki and her infant son to a local motel .	however UNK they are found by their mother UNK and the baby 's mother UNK mrs. UNK UNK is left with no one to help	3	3	3

Model Used	Context Sentence	Generated Sentence	G	R	I
event2sentence	both men are determined to bed the young woman , but they quickly discover patty is more interested in engaging in spirited discussions about the pressing moral and sexual issues of the day than surrendering her virginity to either one of them .	karen and eddie 's latest ended towards the end of the escape breaks in the confusion and patrick 's personal film love between the other time poor boyfriend matters UNK a days later , this time , on the other is the next day after the recording to the boys " present on the island is pair of the dead and dead of them for the next day krishnaveni is	2	1	3
predictEvent	in a mock game of football donald is UNK by the ants .	the next day , donald and his friends go on a date .	5	3	1
seq2seq	the other one hears his screams and goes to investigate .	there UNK the man is heard screaming in the woods .	4	3	1
Gold Standard	kat and nick end up together .	jeffrey , the main reason for all the trouble , learns absolutely nothing .	4	3	2
event2sentence	the commanding general , however , had not forgotten the proposal and was determined to have mulan as a son-in-law and served in the government .	he tells her father is nothing lead to me and word where " prakash catch and his new journey hero sympathetic to his response to morrison accident waiting for his part two will soon meets three of the were both the son is on the next day this is the next morning deeply the boy 's business is angered as events v. insults otto UNK , turns out on the	1	2	2
predictEvent	buford declines because he is in the middle of something important .	buford , buford , and buford go to buford 's house , where buford tells buford that buford has been killed .	5	3	2
seq2seq	the film crew is given behind the scenes access as the noted liberal goes up against his rivals .	the film ends with the characters celebrating their victory .	5	2	2
Gold Standard	barnaby has them arrested on a burglary charge , and the two are sentenced to be UNK in the ducking stool and then banished to UNK .	but barnaby agrees to drop the charges if bo peep will marry him .	4	3	3

Model Used	Context Sentence	Generated Sentence	G	R	I
event2sentence	jimmy snaps over his shoulder “ friend of yours to see you ” and abruptly leaves .	they try to their dismay older mother karen steps in a long before before the events are restored to give the failure to take revenge and manolo and security guards employee it patrick awkward in phil touches the middle school donkey continued asking starting to the rest of the explosion is not wanting to escape from guilt of audiences move in new law college student friends UNK and UNK gestapo	1	1	3
predictEvent	upon returning , nico returned to chicago and joined the chicago police department .	the next day , the police arrive at the house and arrest the police .	5	3	4
Gold Standard	elliot admits that despite her manipulation of him he has come to like satan and regards her as a friend , something she does not object to .	she replies that heaven and hell can be found on earth ; it is up to humans to choose .	5	3	4
event2sentence	all three ’s lives change , when vicky joins karan and sneha in college , and sneha begins feeling drawn towards vicky .	he gains sexual instructions phil personal is younger karl excited while professor henry ’s cat john UNK is a lady james boxing UNK and tokyo , still upset , mrs. jo ’s manager is the lawyer neil suraj is the bank president james escape from a ball written from ivan , beginning to the distant cost them doctor lying silver to him to expand to the remaining action vanessa name	1	1	2
predictEvent	picard , UNK , and several ba ‘ ku are transported onto the son ‘ a ship .	the film ends with the two of them in the middle of the night .	3	1	1
seq2seq	abner is faced with escaping from the footsteps of the people , a bike , an electric tram , and other vehicles .	the next day UNK the team has to go to the UNK to see the city UNK but find the UNK ’ message for the	2	1	3
predictEvent	anna and louis then orchestrate a brilliant deception from their hiding spot in the forest : louis uses his horn to replicate the sound of a bugle charge , as anna “ attacks ” the area with harmless fireworks .	the next morning , the police arrive at the house .	5	2	2

Model Used	Context Sentence	Generated Sentence	G	R	I
seq2seq	naina is a successful photographer and has taught herself that a less than fulfilling life is enough for her .	she finds a life of a woman who had a crush on her and her father UNK who has been UNK by her husband UNK	4	3	3
event2sentence	the movie goes through series of flashbacks of old memories during their college days .	su-kyoung is devastated and unhappy with her to make matters time bhanu his love is a " UNK widow sam UNK works as promising top of shock and dejected raja female students are and van UNK thinks this karan is UNK is a UNK is mrs. jiro 's character shackelford is basically humanity 's mother UNK the real empty childhood empty release from mark audition for midst of ravi fully	1	2	2
Gold Standard	eventually , feeling inferior , munna decides to leave UNK to raj , who can give her a better life than he can .	the matter is n't resolved yet though , as UNK hears of this on the film 's opening night .	3	3	3
predictEvent	the sentry fish sees this and pulls out his own fishing pole with baited hook and the worm jumps onto the hook and the sentry fish reels him in and eats him !	as the credits roll , the camera zooms into the distance .	5	2	3
Gold Standard	leaving fred and george 's new shop , harry , ron and hermione notice draco and narcissa associating with death eaters in UNK and UNK .	harry believes voldemort has made draco a death eater , but ron and hermione are sceptical .	5	3	3
event2sentence	several meteors fall in a field in england .	UNK , farman goes out there , she will be poor in the end of the very change at the age morning the way later , one of the day before night , while the train on the reunion is about the day before afterwards , patrick 's brother is dismayed by the frustration of radio note : as the two escape from the end of the action addresses the	1	1	2
seq2seq	they are chased by keller , after he realizes that they know what happened .	in the course of the film UNK ian and sally are reunited with their mother and are forced to spend their last night together together	3	4	3

Model Used	Context Sentence	Generated Sentence	G	R	I
event2sentence	* the acid house : an acid trip and a bolt of lightning result in amiable UNK coco brice exchanging bodies with the baby of a middle class couple .	one of the children are able to find that his further distance saddened results is unsuccessfully trying to get willing joined tommy , unbeknownst to both elmo 's friends , a former staff catch the doctor burns will infamous and big pete is a frustrated children who will be out of anger , he often wounds her husband returns from the dying , sylvester 's discovery , mason 's entire	1	1	1
predictEvent	he encounters his brother , a hippie living in venice beach , and falls for an attractive flower power hippie girl who has a knack for making pot brownies .	the film ends with the two of them in a car .	5	3	1
seq2seq	the twins escape , run home and promise to help their parents with their work as farmers .	the two friends are reunited .	5	4	3
Gold Standard	he tries to resist her , but he finds himself falling for her as well .	they soon blossom into a passionate love , but they must keep their affair a secret from his wife and the school .	5	5	5
Gold Standard	UNK insists that the UNK are pacifist , but the doctor tests this claim by ordering ian to take UNK , UNK 's love , to the daleks in exchange for the confiscated fluid link .	UNK punches ian to the ground , showing to the UNK that they can fight for some things .	5	5	5
predictEvent	the x-men are able to supply the president with files from stryker 's private offices , and the professor warns him that humans and mutants must work together to build peace or they will destroy each other through war .	however , they are ambushed by the president of the nsa .	5	4	5
event2sentence	this is where bhagat launched his UNK revolution , and became popular with the masses , especially the younger generation , laborers , and farmers , so much so that his popularity UNK that of gandhi himself .	the film is on a distant about a map is not wanting the ship 's content at that humanity 's mandy 's such the weak will are intent on getting on the way back in the end of the robbery is george UNK is n't where his UNK UNK on the way of anton 's corpse is the 1941 that caroline smith robert hart and royal homeless and confused and	1	1	1

Model Used	Context Sentence	Generated Sentence	G	R	I
seq2seq	while in custody for reckless driving , the landlord learns of the baby 's value through the news .	now UNK she is led to a mental asylum .	5	4	5
predictEvent	additionally , one loan shark takes UNK 's own trumpet , and smacks him across the face with it .	the next morning , the shark is found dead in the middle of the night .	5	3	5
seq2seq	by giving the last bit of elemental energy in destroying the stone , he had broken that seal and set teridax free , but vakama just states that " for some reason , he does n't scare me anymore " .	then they both attack the other side of the tower UNK but they are all captured .	3	2	3
event2sentence	after sleeping with vicki , walter reveals his history of molesting little girls .	the team hijacks the land , on the girls are there , the school slow confesses that festival , they both request and delighted at the murder secluded dog escaped pieces of both of his of all of the brother and all of my rather than return troubles for those party ending up on the part of solo 's wreckage behind the film centres on the robbery says their relationship	2	1	1
Gold Standard	channing has a similar ambition : to become a jockey .	her father refuses to let her train , as he believes it 's too dangerous ; he retired from training after channing 's mother , also a jockey , died in a racing accident , several years before .	4	5	5
predictEvent	in the house , mama gives change to an older man , mr. johnson , who has paid for the beer she is holding in her hand .	he is able to find a way to save his life .	5	1	3
Gold Standard	meanwhile , mika has prepared a large cake in the shape of a castle for the UNK cafe 's second anniversary party being held at the cafe tomorrow .	they store the cake in a UNK case unit so that the cake stays fresh until tomorrow .	5	5	2

Model Used	Context Sentence	Generated Sentence	G	R	I
event2sentence	a novel by norwegian writer sigrid boo , vi som gar UNK has an almost identical plot to UNK 's original book .	they meet a doctor on earth , the river 27 aware of mr. kevin , UNK and new UNK and UNK who are most people are more most of years later irons , however , your second son zach is army principal henry 's stern discovery UNK sam 's father-in-law santiago is now curious , the drama became the real bad a particularly bad daughter lady little arthur 's relationship	1	1	5
seq2seq	the main setting is the tolstoy country estate of UNK UNK .	in the present UNK the son of the protagonist UNK UNK UNK UNK is a medical health teacher who has been UNK to the past	3	3	4
predictEvent	the film shows complexities involved in straight , gay , lesbian , and bisexual relationships .	in the end , the film ends with a happy ending .	4	5	1
Gold Standard	keaton stares at the scene , places a ' for sale ' sign with the heap and walks off with UNK .	the new york times movie review said , " one week , a buster keaton work , has more fun in it than most UNK , UNK comedies . "	3	2	4
seq2seq	oki , still under the spell , does not defend puja , but instead suggests that they prepare food for the visitors .	the next day UNK the boy is found dead .	5	5	5
event2sentence	they quickly do and she joins them , and begins to drive the bus away from camp , only to have UNK pop up at the door and try to get in .	her boss is not long after embarking on his way , unknown to the path of three years earlier , in the end of the support role of both the next day though the you teenage circumstances to carry off the theatre union drama justice insult each fbi hits the village owner tommy UNK is a village henry amber UNK and lisa UNK boys tommy UNK 2 spite of the	1	1	3
predictEvent	loren 's friends turn a blind eye to his activities for a while .	the film ends with the two of them walking away together .	5	1	2
seq2seq	ed discovers that the love he left in london has married in his absence , allowing him and paula to be together .	later UNK mark and maggie are married UNK but danny is not alone .	3	2	3

Model Used	Context Sentence	Generated Sentence	G	R	I
event2sentence	upon seeing the creature , she flees in an initial state of panic , but later hears the creature sing and discovers it is not dangerous but has a lovely singing voice .	the film is on a distant about a map is not wanting the ship 's content at that humanity 's mandy 's such the weak will are intent on getting on the way back in the end of the robbery is george UNK is n't where his UNK UNK on the way of anton 's corpse is the 1941 that caroline smith robert hart and royal homeless and confused and	2	1	2
Gold Standard	however , lucy is determined to steal jenna 's position , so she frames jenna .	from this point on , all of jenna 's work to rebuild her life as a thirty year old woman has gone to waste .	5	4	4
Gold Standard	zoe is a bit hesitant at first and she agrees to meet him at the pub that evening .	zoe reveals to her children that the man 's name is david and her children seem excited about this , as there is the allusion to ' david ' beckham .	5	5	3
seq2seq	alice , deeply upset , leaves dan without telling him where she is going .	on the other hand UNK ellen becomes obsessed with the idea of a " perfect UNK " and begins to think she is in love	4	2	3
event2sentence	scotty forces cheryl into the cellar and locks it , but his girlfriend shelly is also possessed by a demon , which spies on her from outside her bedroom as she is changing before hurling itself at her .	kang UNK and UNK , in june 2008 , who is plan was great interest is important as he rides into the film was part of the four spirit is in the very pair of these events are running for the next day each you were , when disturbed , upon request from the rich is a short while thereafter , elias forced to ship and heartbroken , silent suspects	1	1	1
predictEvent	next day , bobby 's things are seen near the river and soma claims that panicker has killed bobby as he came to about panicker 's true color .	the next day , bobby goes to his house and finds that he has gone missing .	4	3	4
predictEvent	he then manages to find it as jigen and goemon head for the caribbean island .	the film ends with the two of them .	4	2	2
seq2seq	sylvester thinks he 's in for a free meal ... until the tide comes in and washes sylvester away .	sylvester then puts down a stream of water UNK and chases the cat out of the hole UNK but it falls down UNK landing down	3	2	3

Model Used	Context Sentence	Generated Sentence	G	R	I
Gold Standard	paul finds an opportunity to meet cristina and eventually reveals how the two of them are connected .	desperately needing one another , they begin to develop a relationship .	5	4	4
event2sentence	after the russian empire is overthrown and the soviet union is born , most of the western diplomats evacuate .	while the others learn from the end of the next day after the entire time sam chasing the UNK trouble between begins to come on his way for revenge for the next ca n't think don norton akash 1994 doubts humanity would no until one of distracted , hank and nineteen years later , the side of the next morning knowing where heartbroken , this , vowing revenge , although	1	1	1
event2sentence	larry digs up the front lawn in search of the water UNK valve .	before the doctor on the morning after , as the woman 's marriage by this time it is new time says it is a far may n't charge " around the time the time count janine UNK , as it is time to be sure how much shot between hope , on the outside the ball ' arguing attempt to achieve this becky is a wealthy rick UNK UNK UNK	2	1	1
seq2seq	cass gets a job as a bouncer outside nightclubs , working for ray , a long-time friend from his UNK days .	their relationship is interrupted by the arrival of a young woman UNK who has been having an affair with the married man .	5	2	4
Gold Standard	still wishing to reconnect with his father , paul sets off to find him .	they first meet up at a petrol station , and later at the scene of an accident .	5	4	4
predictEvent	after an accident destroys her father 's jeep , trishna goes to work for jay , and they fall in love .	the next day , the police arrive at the house .	5	5	4
event2sentence	returning to UNK , isabelle 's mother confronts jenkins in front of the congregation .	the film is on a distant about a map is not wanting the ship 's content at that humanity 's mandy 's such the weak will are intent on getting on the way back in the end of the robbery is george UNK is n't where his UNK UNK on the way of anton 's corpse is the 1941 that caroline smith robert hart and royal homeless and confused and	2	1	1

Model Used	Context Sentence	Generated Sentence	G	R	I
predictEvent	once UNK and kyung-min become fully aware of the situation , they contact UNK who is able to hide the two singers in a secret location .	the film ends with the two of them in the middle of the night .	5	4	3
Gold Standard	zachary 's home will be taken away , and toby gets in trouble with his father , and cal is very upset at him and gives him back all the money he borrowed from him and tells him his mother is never coming back like wayne .	toby gives a country girl he likes a necklace , which was his mother 's , but she gives it back to him .	5	3	4
seq2seq	frank is n't exactly thrilled with his latest assignment , but when his employer turns up the pressure , he has little choice but to deliver .	he is so impressed UNK he agrees to meet with a new client UNK jack UNK UNK who is also a UNK for the american	5	3	3
Gold Standard	slade grabs a chair as a weapon , but amos ducks , and slade falls through an open window to his death .	amos is cleared of all wrongdoing , but the newspaper story attracts the notice of UNK beg , ali khan 's prime minister .	4	4	4
seq2seq	white explains that their UNK contract specifies the content of the album : one original song per side , with the balance consisting of UNK material , but the original songs must be " snappy " , not ballads .	the next day UNK as daryl is walking home UNK a woman named natalie appears in the room UNK and he looks at her as	5	2	3
event2sentence	realising that the rules of chivalry have not been met , sir UNK declares the duel void and declares war .	angry , he brings her mother , who is the dying derek was annie is the middle brother works for a drug , nina is based on the advice judith and pinocchio 's trip to susan 's " UNK and his new lawyer people land rushing destroyed the court bumps off the blue mayor without warning , the money managed to light to the confusion shifts to stop seven of	3	2	2
predictEvent	the interview cuts to kenta and UNK running through the streets of stockholm filmed with a UNK lens .	the film then cuts to the present day , where she is reunited with her husband .	5	5	3