

Movie Plot Analysis via Turning Point Identification

Appendix

Pinelopi Papalampidi Frank Keller Mirella Lapata

Institute for Language, Cognition and Computation

School of Informatics, University of Edinburgh

10 Crichton Street, Edinburgh EH8 9AB

p.papalampidi@sms.ed.ac.uk, {keller,mlap}@inf.ed.ac.uk

1 Model Details

Synopsis Encoder In all tasks we use a synopsis encoder in order to contextualize the sentences in the synopsis. We employ an LSTM network as the synopsis encoder which produces sentence representations h_1, h_2, \dots, h_T , where h_i is the hidden state at time-step i , summarizing all the information of the synopsis up to the i -th sentence. We use a Bidirectional LSTM (BiLSTM) in order to get sentence representations that summarize the information from both directions. A BiLSTM consists of a forward LSTM \vec{f} that reads the synopsis from p_1 to p_N and a backward LSTM \overleftarrow{f} that reads it from p_N to p_1 . We obtain the final representation cp_i for a given synopsis sentence p_i by concatenating the representations from both directions, $cp_i = h_i = [\vec{h}_i; \overleftarrow{h}_i]$, $h_i \in \mathbb{R}^{2S}$, where S denotes the size of each LSTM.

Entity-Specific Encoder This encoder is used to evaluate the contribution of entity-specific information to the performance of our models. We use a word embedding layer to project words w_1, w_2, \dots, w_T of the i^{th} synopsis sentence p_i to a continuous vector space \mathbb{R}^E , where E the size of the embedding layer. This layer is initialized with pre-trained entity embeddings. Next, we use a BiLSTM as described in the case of the synopsis encoder. On top of the LSTM, we add an attention mechanism, which assigns a weight a_i to each word representation h_i . We compute the entity-specific representation pe_i of the i^{th} plot sentence as the weighted sum of word representations:

$$e_j = \tanh(W_h h_j + b_h), \quad e_j \in [-1, 1] \quad (1)$$

$$a_j = \frac{\exp(e_j)}{\sum_{t=1}^T \exp(e_t)}, \quad \sum_{j=1}^T a_j = 1 \quad (2)$$

$$pe_i = \sum_{j=1}^T a_j h_j, \quad e \in \mathbb{R}^{2S} \quad (3)$$

where W_h and b_h are the attention layer's weights.

2 Implementation Details

Pre-trained Sentence Encoder The performance of our models depends on the initial sentence representations. We experimented with using the large BERT model (Devlin et al., 2018) and the Universal Sentence Encoder (USE) (Cer et al., 2018) as pre-trained sentence encoders in all tasks. Intuitively, we expect USE to be more suitable, since it was trained in textual similarity tasks which are more relevant to ours. Experiments on the development set confirmed our intuition. Specifically, on the screenplay TP prediction task, annotation distance D dropped from 17.00% to 10.04% when employing USE instead of the BERT embeddings in the CAM version of our architecture.

Hyper-parameters We used the Adam algorithm (Kingma and Ba, 2014) for optimizing our networks. After experimentation, we chose an LSTM with 32 neurons (64 for the BiLSTM) for the synopsis encoder in the first task and one with 64 neurons for the encoder in the second task. For the context interaction layer, the window l was set to two sentences for the first task and 20% of the screenplay length for the second task. For the entity encoder, an embedding layer of size 300 was initialized with the Wikipedia2Vec pre-trained word embeddings (Yamada et al., 2018) and remained frozen during training. The LSTM of the encoder had 32 and 64 neurons for the first and second tasks, respectively. Finally, we also added a dropout of 0.2. For developing our models we used PyTorch (Paszke et al., 2017).

Data Augmentation We used multiple annotations for training for movies where these were available and considered reliable. The reasons for this are twofold. Firstly, this allowed us to take into account the subjective nature of the task dur-

Goldstandard
<ul style="list-style-type: none"> • Sixteen-year-old Minnesota high-schooler Juno MacGuff discovers she is pregnant with a child fathered by her friend and longtime admirer, Paulie Bleeker. • All of this decides her against abortion, and she decides to give the baby up for adoption. • With Mac, Juno meets the couple, Mark and Vanessa Loring (Jason Bateman and Jennifer Garner), in their expensive home and agrees to a closed adoption. • Juno watches the Loring marriage fall apart, then drives away and breaks down in tears by the side of the road. • Vanessa comes to the hospital where she joyfully claims the newborn boy as a single adoptive mother.
TAM (+ TP views)
<ul style="list-style-type: none"> • Going to a local clinic run by a women’s group, she encounters outside a school mate who is holding a rather pathetic one-person Pro-Life vigil. • With Mac, Juno meets the couple, Mark and Vanessa Loring (Jason Bateman and Jennifer Garner), in their expensive home and agrees to a closed adoption. • Juno and Leah happen to see Vanessa in a shopping mall being completely at ease with a child, and Juno encourages Vanessa to talk to her baby in the womb, where it obligingly kicks for her. • Juno watches the Loring marriage fall apart, then drives away and breaks down in tears by the side of the road. • The film ends in the summertime with Juno and Paulie playing guitar and singing together, followed by a kiss.
Distribution baseline
<ul style="list-style-type: none"> • Once inside, however, Juno is alienated by the clinic staff’s authoritarian and bureaucratic attitudes. • Juno visits Mark a few times, with whom she shares tastes in punk rock and horror films. • Not long before her baby is due, Juno is again visiting Mark when their interaction becomes emotional. • Juno then tells Paulie she loves him, and Paulie’s actions make it clear her feelings are very much reciprocated. • Vanessa comes to the hospital where she joyfully claims the newborn boy as a single adoptive mother.

Table 1: Highlights for the movie "Juno": goldstandard annotations and predicted TPs for TAM (+ TP views) and distribution baseline.

ing training; and secondly, it increased the size of our dataset, which contains a limited number of movies. Specifically, we added triplicate annotations for 17 movies and duplicate annotations for 5 movies.

3 Example Output: TP Identification in Synopses

As mentioned in Section 6, we also conducted a human evaluation experiment, where highlights were extracted by combining the five sentences labeled as TPs the synopsis. In Tables 1, 2, and 3, we present the highlights presented to the AMT workers for the movies "Juno", "Panic Room", and "The Shining", respectively. For each movie we show the goldstandard annotations alongside with the predicted TPs for TAM (+ TP views) and the distribution baseline, which is the strongest performing baseline with respect to the automatic evaluation results.

Overall, we observe that goldstandard highlights describe the plotline of the movie, contain a first introductory sentence, some major and intense events, and a last sentence that describes the ending of the story.

The distribution baseline is able to predict a few goldstandard TPs by only considering the relative position of the sentences in the synopsis. This observation validates the screenwriting theory: TPs, or more generally important events that

determine the progression of the plot, are consistently distributed in specific parts of a movie. However, when the distribution baseline cannot predict the exact TP sentence, it might select one that describes irrelevant events of minor importance (e.g., TP4 for "Panic Room" is a detail about a secondary character instead of a major setback and highly intense event in the movie).

Finally, our own model seems to be able to predict some goldstandard TP sentences, as demonstrated during the automatic evaluation. However, we also observe here that even when it does not select the goldstandard TPs, the predicted ones describe important events in the movie that have some desired characteristics. In particular, for the movie "Juno" the climax (TP5) is the moment of resolution, where Vanessa decides to adopt the baby after all the setbacks and obstacles. Even though our model does not predict this sentence, it does select one that reveals information about the ending of the movie. An other such example is the movie "Panic Room", where the point of no return (TP3) is not correctly predicted, but the selected sentence refers to the same event.

References

Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar,

Goldstandard
<ul style="list-style-type: none"> • On the night the two move into the home, it is broken into by Junior, the previous owner’s grandson; Burnham, an employee of the residence’s security company; and Raoul, a ski mask-wearing gunman recruited by Junior. • Before the three can reach them, Meg and Sarah run into the panic room and close the door behind them, only to find that the burglars have disabled the telephone. • To make matters worse, Sarah, who has diabetes, suffers a seizure. • Sensing the potential danger to her daughter, Meg lies to the officers and they leave. • After a badly injured Stephen shoots at Raoul and misses, Raoul disables him and prepares to kill Meg with the sledgehammer, but Burnham, upon hearing Sarah’s screams of pain, returns to the house and shoots Raoul dead, stating, “You’ll be okay now”, to Meg and her daughter before leaving.
TAM (+ TP views)
<ul style="list-style-type: none"> • On the night the two move into the home, it is broken into by Junior, the previous owner’s grandson; Burnham, an employee of the residence’s security company; and Raoul, a ski mask-wearing gunman recruited by Junior. • Before the three can reach them, Meg and Sarah run into the panic room and close the door behind them, only to find that the burglars have disabled the telephone. • Her emergency glucagon syringe is in a refrigerator outside the panic room. • As Meg throws the syringe into the panic room, Burnham frantically locks himself, Raoul, and Sarah inside, crushing Raoul’s hand in the sliding steel door. • After a badly injured Stephen shoots at Raoul and misses, Raoul disables him and prepares to kill Meg with the sledgehammer, but Burnham, upon hearing Sarah’s screams of pain, returns to the house and shoots Raoul dead, stating, “You’ll be okay now”, to Meg and her daughter before leaving.
Distribution baseline
<ul style="list-style-type: none"> • On the night the two move into the home, it is broken into by Junior, the previous owner’s grandson; Burnham, an employee of the residence’s security company; and Raoul, a ski mask-wearing gunman recruited by Junior. • Unable to seal the vents, Meg ignites the gas while she and Sarah cover themselves with fireproof blankets, causing an explosion which vents into the room outside and causes a fire, injuring Junior. • To make matters worse, Sarah, who has diabetes, suffers a seizure. • While doing so, he tells Sarah he did not want this, and the only reason he agreed to participate was to give his own child a better life. • As the robbers attempt to leave, using Sarah as a hostage, Meg hits Raoul with a sledgehammer and Burnham flees.

Table 2: Highlights for the movie “Panic Room”: goldstandard annotations and the predicted TPs for TAM (+ TP views) and distribution baseline.

Goldstandard
<ul style="list-style-type: none"> • Manager Stuart Ullman warns him that a previous caretaker developed cabin fever and killed his family and himself. • Hallorann tells Danny that the hotel itself has a “shine” to it along with many memories, not all of which are good. • After she awakens him, he says he dreamed that he had killed her and Danny. • Jack begins to chop through the door leading to his family’s living quarters with a fire axe. • Wendy and Danny escape in Hallorann’s snowcat, while Jack freezes to death in the hedge maze.
TAM (+TP views)
<ul style="list-style-type: none"> • Jack’s wife, Wendy, tells a visiting doctor that Danny has an imaginary friend named Tony, and that Jack has given up drinking because he had hurt Danny’s arm following a binge. • Hallorann tells Danny that the hotel itself has a “shine” to it along with many memories, not all of which are good. • Danny starts calling out “redrum” frantically and goes into a trance, now referring to himself as “Tony”. • When Wendy sees this in the bedroom mirror, the letters spell out “MURDER”. • Wendy and Danny escape in Hallorann’s snowcat, while Jack freezes to death in the hedge maze.
Distribution baseline
<ul style="list-style-type: none"> • Jack’s wife, Wendy, tells a visiting doctor that Danny has an imaginary friend named Tony, and that Jack has given up drinking because he had hurt Danny’s arm following a binge. • Jack, increasingly frustrated, starts acting strangely and becomes prone to violent outbursts. • Jack investigates Room 237, where he encounters the ghost of a dead woman, but tells Wendy he saw nothing. • When Wendy sees this in the bedroom mirror, the letters spell out “MURDER”. • He kills Hallorann in the lobby and pursues Danny into the hedge maze.

Table 3: Highlights for the movie “The Shining”: goldstandard annotations and the predicted TPs TAM (+ TP views) and distribution baseline.

et al. 2018. Universal sentence encoder. <i>arXiv preprint arXiv:1803.11175</i> .	Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. <i>arXiv preprint arXiv:1412.6980</i> .
Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. <i>arXiv preprint arXiv:1810.04805</i> .	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.

Ikuya Yamada, Akari Asai, Hiroyuki Shindo, Hideaki Takeda, and Yoshiyasu Takefuji. 2018. Wikipedia2vec: An optimized tool for learning embeddings of words and entities from wikipedia. *arXiv preprint 1812.06280*.