A Supplement Material

A detailed visualization of our model, described in Section 3 of the main paper is shown in Fig. 10.

A.1 Dataset Details

We train and test our model on the *Modeling Naive Psychology of Characters in Simple Commonsense Stories* dataset (Rashkin et al., 2018). It contains narrative stories where each sentence is annotated with a character and a set of human need categories from two inventories: Maslow's (with five coarse-grained) and Reiss's (with 19 fine-grained) categories. Figure 6 portraits the labels in Reiss and Maslow and their relation. Figures 7 and 8 depict the data distribution for the training and dev set for Reiss and Maslow respectively. As in prior work we select the annotations that display the "majority label" i.e., categories voted on by ≥ 2 workers. Since no training data is available, similar to prior work we use a portion of the devset as training data, by performing a random split, using 80% of the data to train the classifier, and 20% to tune parameters. We use ConceptNet version 5.6.0 to extract commonsense knowledge.

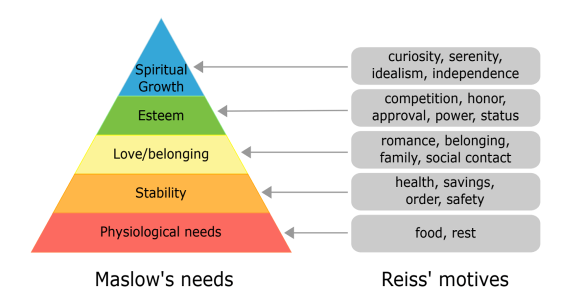


Figure 6: Maslow and Reiss Labels

A.2 Training Details

In training, we minimize the weighted binary cross entropy loss to train our multi-label classifier with the Adam optimizer (Kingma and Ba, 2014) with an initial learning rate of 0.001, a dropout-rate of 0.5 (dropout is applied to the input of each LSTM layer) and batch size of 32. We use 300 dimensional word embeddings and a hidden size of 100 for all Dense Layer and k = 3 for the selection of top ranked paths. For Maslow labels, we use L2 regularization with $\lambda = 0.01$, For Reiss labels, we use L2 regularization with $\lambda = 0.1$.

A.3 Concept to Human Needs

We manually aligned the human need categories to concepts in ConceptNet. We used the name of the human needs to map them to identically named concepts from ConceptNet, except for 3 human needs classes, which are as follows (Table 6):

Concepts	**Human needs**
tranquility	safety
serenity	calm
contact	social

Table 6: Concepts corresponding to Human needs

For Maslow's labels we use the mapping for Reiss, as Maslow's categories are a subset of the Reiss categories, as shown in Figure 6.

A.4 Human evaluation

We conduct human evaluation to test the effectiveness and relevance of the extracted commonsense knowledge paths. We randomly selected 50 sentence-context pairs with their gold labels from the dev set and extracted knowledge paths that contain the gold label (using CC+PPR for ranking). We asked three expert evaluators to decide whether the paths provide relevant information about the missing links between the concepts in the sentence and the human need (gold label). We asked them to assign scores according to the following definitions:

+2: the path specifies perfectly relevant information to provide the missing link between the concepts in the sentence and the human need.

+1: the path contains a sub-path that specifies relevant information to provide the missing links between the concepts in the sentence and the human need.

0: when the path is irrelevant but the starting and the ending nodes stand in a relation that is relevant to link the sentence and the expressed human need. (In this case, either the path selected by our algorithm is not relevant or there is no relevant path connecting the nodes given the context.)

-1: the path is completely irrelevant.

Figure 9 depicts the distribution of assigned scores (based on the majority class). It shows that

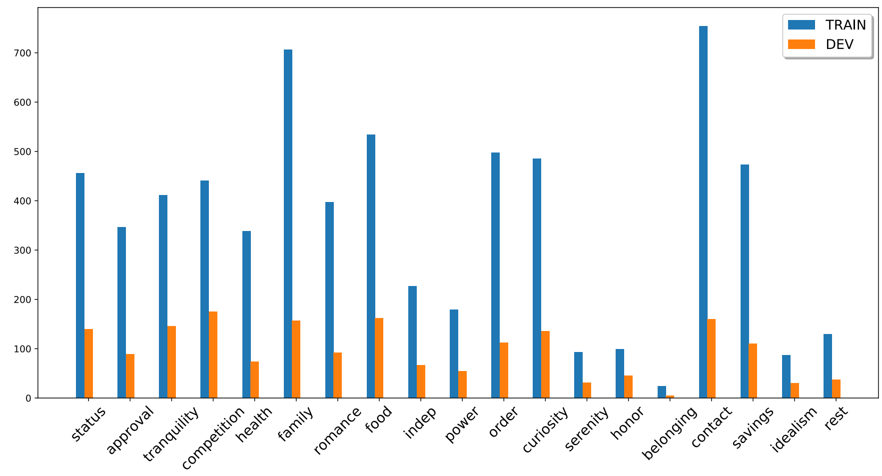


Figure 7: Train and Dev data statistics for Reiss Classification.

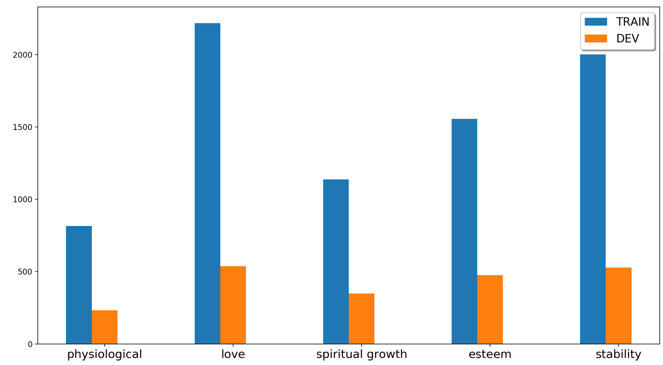


Figure 8: Train and Dev data statistics for Maslow Classification.

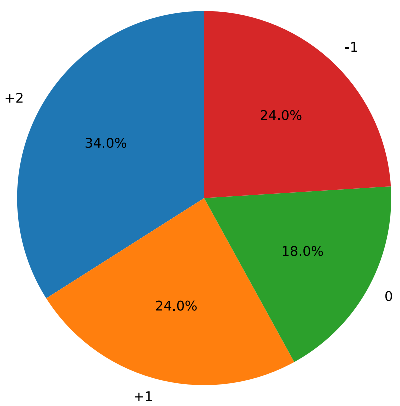


Figure 9: Human evaluation: Distribution of scores.

in 34% of the cases our algorithm was able to select a relevant commonsense path. In another 24% of cases a sub-part of the selected path was still considered relevant.

A.5 Model Analysis and Visualization

We study the visualization of attention distributions produced by our model. We provide examples for different scenarios. Here we show the results found by our best model i.e., BiLSTM+Self-Attention+Gated-Knowledge with CC+PPR as path selection method.

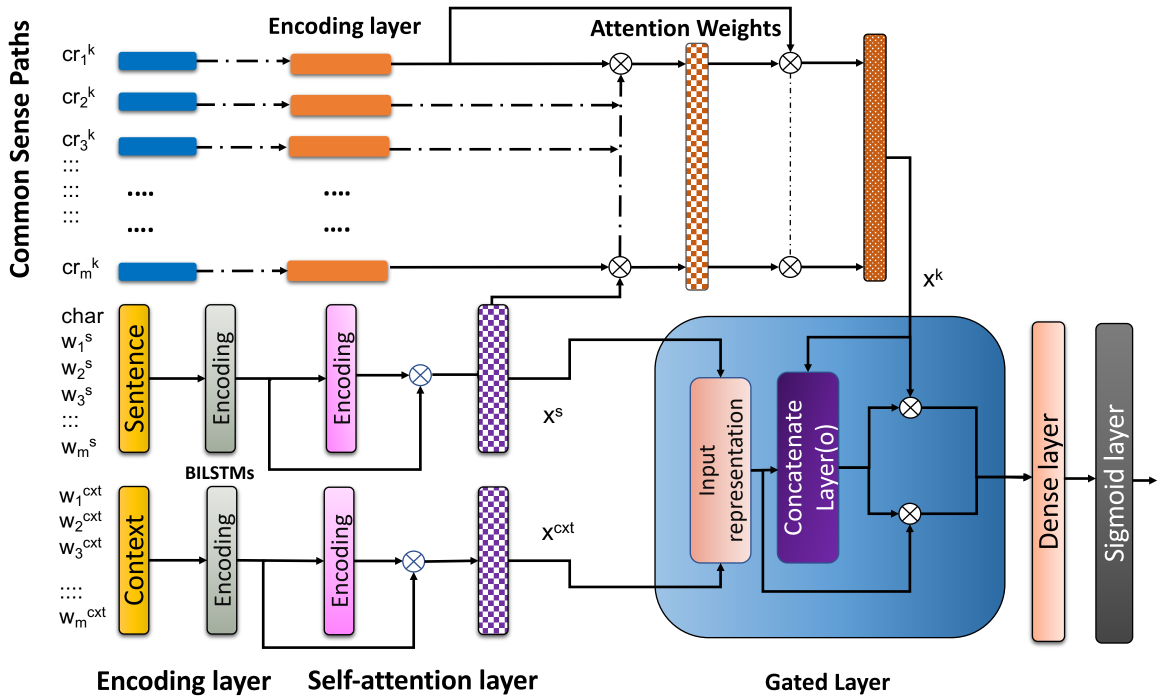


Figure 10: Full model

Case 1: Inclusion of knowledge path improves the performance when there is no context.

Context: No Context
Sentence: Tina was out for a walk in the street.
True Label: *Health*
Predicted without Knowledge: *Serenity*
Predicted with Knowledge : *Health*

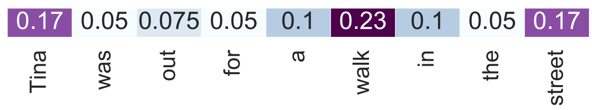


Figure 11: Example 1: Visualizing the attention weights of the input sentence and of selected commonsense paths.

Case 2: Inclusion of knowledge paths improves the precision of the model.

Context: No Context
Sentence: Noah wanted to play golf against Nick.
True Label: *Competition*
Predicted without Knowledge: *Competition, Curiosity*
Predicted with Knowledge : *Competition*

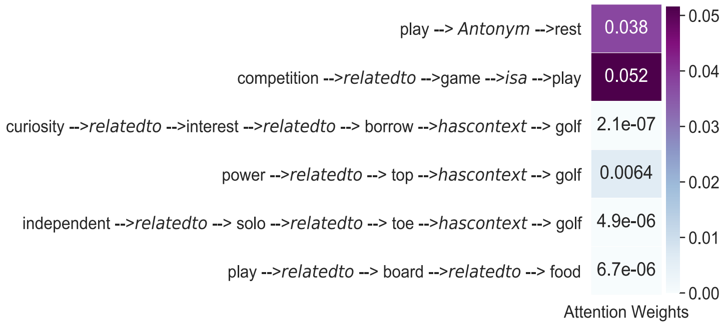
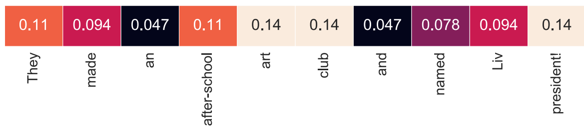


Figure 12: Example 2: Visualizing the attention weights of the input sentence and of selected commonsense paths.

Case 3: Inclusion of knowledge paths improves the recall of the model

Context: Liv was a budding artist and she loved painting. She wanted to go to art classes, but her school didn't offer any!, So Liv got together with her friends and began brainstorming. They decided to form their own art group at the high school.
Sentence: They made an after-school art club and named Liv president!
True Label: *Independent, Curiosity, Contact*
Predicted without Knowledge: *Contact*
Predicted with Knowledge : *Independent, Curiosity, Contact*



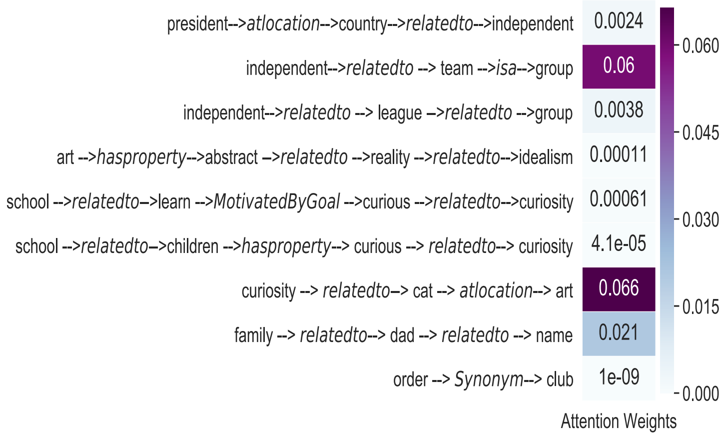
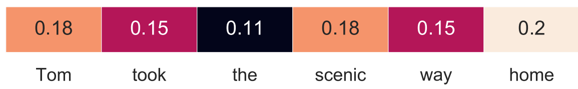


Figure 13: Example 3: Visualizing the attention weights of the input sentence and of selected commonsense paths.

Case 4: In this case our model fails to attend to the relevant path. Although the graph-based ranking and selection algorithm were able to extract a relevant knowledge path, the neural model fails to correctly pick (attend to) the correct path.

Context: Tom was driving his car. He wanted to take a scenic way home. He deliberately passed his exit. Tom saw many beautiful trees.
Sentence: Tom took the scenic way home.
True Label: *Serenity*
Predicted without Knowledge: *Independent, Curiosity*
Predicted with Knowledge : *Family, Independent, Curiosity, Serenity*



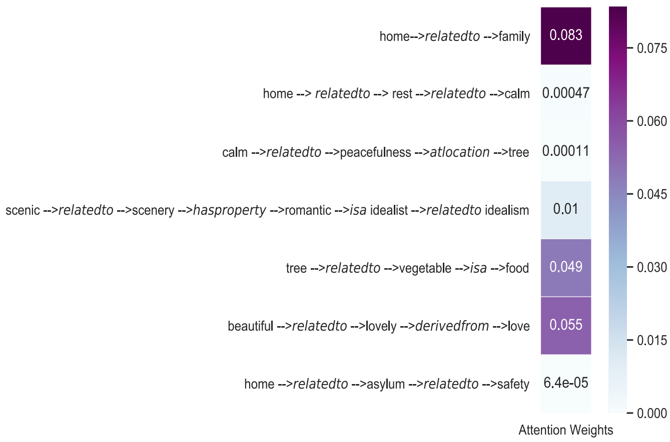


Figure 14: Example 4: Visualizing the attention weights of the input sentence and of selected commonsense paths.