

# Better Conversations by Modeling, Filtering, and Optimizing for Coherence and Diversity: Supplemental Material

## A Determining Cut-off Coherence Score

As shown in Figure 5, the scores on the training set roughly follow a normal distribution with a slight tail on the negatively correlated side. We make the assumption that the data fit a normal distribution and estimate parameters  $\mu = 0.25$  and  $\sigma = 0.22$ . We set the cut-off to  $+2\sigma$  so that it accounts for 95% of the scores and does not severely filter the number of resulting examples in the dataset.

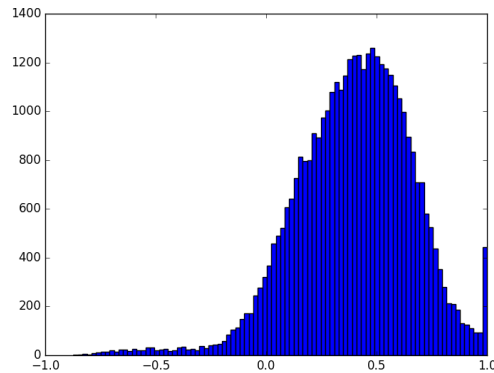


Figure 5: Histogram of coherence scores on the training set before filtering.

## B Dialogue Examples from the Training Set

Tables 4-7 show dialogue examples from the training set with different coherence scores. The instances in Table 4 are strongly lexically coherent by repeating the same words in the response when coherence scores are over 0.9. When scores are between 0.7 and 0.9 (Table 5), the responses use synonyms to words in the context but still remain semantically coherent. When scores are between 0.4 and 0.6 (Table 6), the responses are coherent but very dull; they are getting incoherent when scores are lower (Table 7).

Dialogue contexts	Responses
yeah? ## you're new in town, right?	you call this a town?
so you can have marie all to yourself? ## you're so selfish, catherine.	you've always been selfish.
wait. ## where are you going? ## to find the president.	the president, he lives!
that's not good! ## it 's fine. it itches a bit at first, but then it stops.	of course it itches!
you stay out of this. ## so you want water, huh? ## that's right.	well, i got water.

Table 4: Dialogue examples with coherence score  $C(\mathbf{x}, \mathbf{y}) > 0.9$  (“##” in the context denotes turn boundaries).

Dialogue contexts	Responses
not quite yet. ## call your grandfather to pick me up. ## i want to go home.	grandpa 's not here .
some kind of whisky nobody's ever heard of. ## why don't you bring your own bottle?	give him the best bourbon you got, hot stuff, and don't be gone too long.
put your head on my shoulder? ## denny ## i just want to remember.	i don't think my neck even bends, anymore.
i don't even know where to start. ## kitchen. ## definitely the kitchen.	specifically the stove.
the problem is, the liquid just stays in your gut. ## i don't know what to do.	well, obviously it's not getting absorbed into the bloodstream.

Table 5: Dialogue examples with coherence score  $0.7 < C(\mathbf{x}) < 0.9$  (“##” in the context denotes turn boundaries).

Dialogue contexts	Responses
i'm gonna hold it. ## take a look at it. ## make sure it was an accident.	doesn't sound right.
in fact, thank you for underplaying it. ## so the boy becomes a man	it's amazing.
oh, sting? ## little bit. ## can we stop at the drug store?	oh, uh, don't worry.
there's no one in there! ## it's not gonna happen. ## it's a private party.	nice going .
give me that pad. ## what are you gonna do? ## just watch the old mastermind.	what are you doing?

Table 6: Dialogue examples with coherence score  $0.4 < C(\mathbf{x}, \mathbf{y}) < 0.6$  (“##” in the context denotes turn boundaries).

Dialogue contexts	Responses
otherwise it's all over do whatever he wants ## alright. ## so?	how much do you want?
i'll let it go. ## don 't worry. ## i've totally let it go.	it's all water under the bridge.
i'm so sorry. ## mag? ## thank you.	i'm going to go for a walk.
i dont want it. ## why ## ok	you want the car.
he's thinking about leaving the theatre. ## we've saved you a seat here.	come on. look at her hair.

Table 7: Dialogue examples with coherence score  $C(\mathbf{x}, \mathbf{y}) < 0.4$  (“##” in the context denotes turn boundaries).