Appendix

A Examples from Qualitative Analysis

In Table 8, we present cherry-picked examples for the error patterns introduced in Section 5.6.

Imaginary words. Operating on a sub-word level, the seq2seq model is capable of producing nonexistent words by concatenating unrelated WordPieces. Indeed, we have encountered examples of such made-up words in the outputs of the seq2seq model on all 4 tasks. Typically, this happens when the model encounters a rare word in the input text. LASERTAGGER, which is trained on the word level, is immune to this specific type of error.

Premature end-of-sentence. A seq2seq model produces sequences of arbitrary length by generating the end-of-sentence (`EOS`). We have seen the seq2seq model generate `EOS` prematurely, resulting in an abrupt sentence ending. In extreme cases, and especially on the abstractive summarization task, the model generated `EOS` at the start, effectively producing an empty output. For LASERTAGGER, this type of error is technically possible but very unlikely, and we have not seen it in practice. (The tagger would need to generate a long sequence of `DELETE` tags—something it has not seen in the training data.)

Repeated phrases. Another type of error specific to the seq2seq model is repetition of information—either single words or entire phrases. In the sentence splitting task, the seq2seq model would often repeat parts of the sentence twice (before and after the splitting symbol). In the grammar correction and summarization tasks, the seq2seq model would often replace a rare word with another word from the input sentence, thus repeating that word twice. LASERTAGGER can only add words or phrases from its limited vocabulary, which is unlikely to cause repetition.

We observed that the seq2seq model was especially likely to repeat large fragments in the sentence splitting task in cases when there is no obvious good way to split the sentence. Interestingly, in most of these cases LASERTAGGER did not split the sentence at all, by not inserting the sentence-splitting symbol ⟨::::⟩, even though such examples were not present in its training data. In other cases, LASERTAGGER produced a "lazy split" (discussed below).

Hallucination is a known problem for neural networks, but LASERTAGGER is susceptible to it to a much lesser degree. LASERTAGGER can "hallucinate" only by inserting an unexpected word or a short phrase from its vocabulary. We have seen such insertions in the tagger output that rendered the sentence ungrammatical, or simply odd. A seq2seq model is more likely to produce subtly misleading hallucinations that misrepresent the input text, while looking fluent and credible. We have seen examples of the seq2seq model changing the factual details, which is a more dangerous error to make in some scenarios.

Coreference problems. This type of errors is often made by both seq2seq and tagger models. In the most typical instance, a model inserts an incorrect pronoun. In other cases, the model makes an incorrect coreference resolution, by inserting the wrong proper noun. The seq2seq model is more susceptible to the second type of error because of its ability to copy proper nouns from the input sentence. LASERTAGGER will typically not attempt to resolve the coreference and just leave the pronoun intact.

Misleading rephrasing through deletion. We want to highlight that, although LASERTAGGER has a limited set of operations and cannot insert arbitrary phrases, it is not immune to semantic errors and misrepresenting the input text. We have seen examples where deletion of word spans was sufficient to completely alter the meaning of the sentence. A seq2seq model is naturally also prone to this type of errors.

"Lazy splitting." This type of error is specific to LASERTAGGER and the sentence splitting task. It occurs when the input sentence is split arbitrarily in the middle, without any modifications made to the resulting two parts. Such splits were usually made at a reasonable point in the input sentence, i.e. yielding two valid grammatical clauses, at least one of which, however, was not a complete sentence.

Imaginary words introduced by a seq2model (sentence splitting task).
input: Zenica (Cyrillic : " Зеница ") is an industrial city (the third largest , after Sarajevo and Banja Luka) in Bosnia and Herzegovina and the capital of the Zenica - Doboj Canton of the Federation of Bosnia and Herzegovina entity .
output: Zenica (Cyrillic : "**gratulation**еница ") is an industrial city (the third largest , after Sarajevo and Banja Luka) in Bosnia and Herzegovina . ⟨::::⟩ It is the capital of the Zenica - Doboj Canton of the Federation of Bosnia and Herzegovina entity .

Imaginary words indroduced by a seq2model (sentence fusion task).
input: Carboxysomes are found in lithoautotrophically and mixotrophically grown cells . Carboxysomes aid carbon fixation .
output: Carboxysomes , which aid carbon **fixotrophically** and mixotrophically grown cells .

Abrupt sentence end by the seq2seq model (grammar correction task).
input: By the way, my favorite football team is Manchester United, they are brilliant, they have an amazing football players, and they are awesome.
output: By the way, my favorite football team is.

Abrupt sentence end by the seq2seq model (summarization task).
input: It is, I suppose, likely that people with little money must think of money more than the well-off do and that the hungry will dwell on thoughts of food and the thirsty on drink.
output: **It'**

Redundant repetition by the seq2seq model (sentence splitting task).
input: Bethany Noel Mota (born November 7 , 1995) is an American video blogger from Los Banos , California .
output: Bethany Noel **Mota (born November 7 , 1995) is an American video blogger from Los Banos , California .** ⟨::::⟩ **Mota is an American video blogger from Los Banos , California .**

Redundant repetition by the seq2seq model (sentence fusion task).
input: General Fazlollah Zahedi became prime minister . General Fazlollah Zahedi led the military coup .
output: General Fazlollah Zahedi , who led the military coup , **became prime minister** and **became prime minister** .

Redundant repetition by the seq2seq model (grammar correction task).
input: I'm your employee, to serve on your company.
output: I'm your **company**, to serve on your **company**. .

Hallucinations produced by the tagger (grammar correction task).
input: If you are interesedt, find me at 8 clock near the cinema entrance.
output: If you are **play**, find me at 8 clock near the cinema entrance.

Hallucinations produced by the tagger (sentence fusion task).
input: He carried a bow . He was Nerva .
output: **Receiving** a bow , he was Nerva .

Hallucinations produced by the seq2seq model (sentence splitting task).
input: The native block size is 8192 bytes ; inodes are **512** bytes on disk .
output: The native block size is 8192 bytes . ⟨::::⟩ Inodes are **256** bytes on disk .

Hallucinations produced by the seq2seq model (sentence fusion task).
input: Tobacco smokers may also experience more powerful effects with psilocybin . Tobacco smoke exposure decreases the activity of MAO in the brain and peripheral organs .
output: **anthropology** smokers may also experience more powerful effects with psilocybin .

Coreference error made by the seq2seq model (sentence fusion task).
input: He tries to help Yuushi with catching Sayuki 's eye , and encourages the idea of a baseball club .
output: He tries to help Yuushi with catching Sayuki 's eye . ⟨::::⟩ **Yuushi** encourages the idea of a baseball club .

Pronominalization error made by the tagger (sentence fusion task).
input: Sharmistha Mukherjee is an Indian Kathak dancer , choreographer and politician of the Indian National Congress .
output: Sharmistha Mukherjee is an Indian Kathak dancer , choreographer . ⟨::::⟩ And politician . ⟨::::⟩ **He** is of the Indian National Congress .

"Lazy splitting" made by the tagger (sentence splitting task).
input: Home world of the Marglotta located in the Sagittarius Arm .
output: Home world of the Marglotta . ⟨::::⟩ Located in the Sagittarius Arm .

"Lazy splitting" made by the tagger (sentence splitting task).
input: Jesse has a son named Sahvere and a duaghter named Jaylen .
output: Jesse has a son named Sahvere . ⟨::::⟩ And a duaghter named Jaylen .

Misleading rephrasing through deletion, made by the tagger (sentence splitting task).
input: She has an ex-lover , Winnie Mann , with whom she has a son , Wilson and an adopted a Chinese girl , Jang Yin .
output: She has an ex-lover , Winnie Mann , a son , Wilson . ⟨::::⟩ **She was** adopted Chinese girl , Jang Yin .

Misleading rephrasing through deletion, made by the tagger (sentence splitting task).
input: The article proudly notes that the postal service **was in no way responsible** for the 1996 crash of Valujet Flight 592 in the Florida Everglades.
output: The article notes postal service **was responsible** for the 1996 crash of Valujet Flight 592.

Table 8: Illustration of the typical errors produced by the models (cherry-picked examples).