Linguistics and Artificial Intelligence

Assessing the Implicit Syntactic Abilities of the BERT Neural Transformer Model

Mariann Lengyel
R0977681

Abstract

This report details a small-scale experiment investigating the implicit syntactic abilities of the BERT neural transformer model through masked language modelling. The experiment, which is conducted in Hungarian, focuses on sentence structures where the presence of an attractor noun would confuse the model in predicting the correct output, if it did not have an understanding of the hierarchical structure of the sentence and longer-distance dependencies. The report discusses the results of experimenting on a set of 50 sentences where different constituent elements were incorporated around the examined noun-attractor-verb structure. I then conclude by summarizing the most important findings as well as outlining some potential areas for future experimentation.

Methodology

To conduct the experiment, I started by writing a Python script (included in my submission) that relied on the Hugging Face Transformers library and incorporated a model named "SZTAKI-HLT/hubert-base-cc". The model, in whose name the word "hubert" stands for "Hungarian BERT" model, is a specific variant of BERT models that had been trained on the Hungarian language. To assemble the script, I have relied on the class materials, more specifically the notebook of the fourth exercise session. The script takes a Hungarian sentence as an input with a masked token (denoted "[MASK]") and prints out the top predictions that the model suggests for the masked token based on the sentence.

After writing and testing my script, I designed a test set to test the implicit syntactical understanding of the machine. In my set, I have included ten of subsets of five sentences, resulting in a dataset of 50 sentences. Out of the five sentences of each subset, I have written three with a singular main subject and two with a plural main subject. The pattern that defines the subsets is the following:

1. Noun phrase with a relative clause.

2-3. Noun phrases with a relative clause with an adverb modifying the masked verb.

4. Noun phrase with a relative clause, followed by an adverb modifying the masked verb, followed by a direct object phrase.

5. Noun phrase with a relative clause containing one embedded relative clause (which points to the subject of the phrase), followed by direct object phrase.

6. Noun phrase with a relative clause containing one embedded relative clause (which points to the subject of the phrase), followed by an adverb modifying the masked verb, and a direct object phrase.

7. Noun phrase with a relative clause with one embedded relative clause (which points to the object of the phrase), followed by a direct object phrase.

8.. Noun phrase with a relative clause with one embedded relative clause (which points to the object of the phrase), followed by an adverb modifying the masked verb, with direct object phrase.

9. Noun phrase with a relative clause containing two embedded relative clauses (all pointing to the subject of the phrase), followed by an adverb modifying the masked verb followed by a direct object phrase.

10. Noun phrase with a relative clause containing three embedded relative clauses (all pointing to the subject of the phrase), followed by an adverb modifying the masked verb followed by a direct object phrase.

Examples for all these structures are featured in the next section. Due to its extensive nature, I have decided to include the complete test set as a separate file in my submission rather than appending it to the end of this report. Finally, note that proper Hungarian grammar would require the use of a commas for disambiguation. To preserve neutrality, I have removed all commas from all sentences.

Discussion of the results

For demonstration purposes, in the table below I present the first sentence in each subset along with the results. The sentence column contains the input sentences with a masked verb, as well as their English translation. Please note for the second subset that in Hungarian the verb is sometimes implicitly expressed in the adjective (e.g. instead of “the doctors are clever” we say “the doctors clevers” with the word “clever” changing according to the conjugation of “are”), this is the reason for what looks like a double mask in the English translation.

The table also contains two columns for the form of the expected output verb as well as the form of the actual output verb. Finally, I have also included a column for the overall accuracy of the top five guesses to assess how “sure” the model was in its choice of a verb form.

Subset No.	Sentence	Expected VF	Actual VF	Top 5 accuracy
1.	Az orvos akit várnak a betegek [MASK]. *The doctor whom the patients are waiting for [MASK].*	S/3	period	0/5
2.	Az orvos akit várnak a betegek nagyon [MASK]. *The doctor whom the patiens are waiting for [MASK] very [MASK]*	S/3	S/3	5/2
3.	Az orvos akit várnak a betegek itt [MASK]. *The doctors whom the patients are waiting for [MASK] here.*	S/3	S/3	5/5
4.	Az orvos akit várnak a betegek nagyon [MASK] az almát és a körtét. *The doctor for whom the patients are waiting for really [MASK] apples and pears.*	S/3	S/3	4/5
5.	Az orvos akit várnak a betegek mert sietnek [MASK] az almát és a körtét. *The doctor whom the patients are waiting for because they are in a hurry [MASK] apples and pears.*	S/3	S/3	1/5
6.	Az orvos akit várnak a betegek mert sietnek nagyon [MASK] az almát és a körtét. *The doctor whom the patients are waiting for because they are in a hurry really [MASK] apples and pears.*	S/3	S/3	5/5
7.	Az orvos akit várnak a betegek mert szükséges [MASK] az almát és a körtét. *The doctor whom the patients are waiting for because he is needed [MASK] apples and pears.*	S/3	infinitive	2/5
8.	Az orvos akit várnak a betegek mert szükséges nagyon [MASK] az almát és a körtét. *The doctor whom the patients are waiting for because he is needed really [MASK] apples and pears.*	S/3	S/3	2/5
9.	Az orvos akit várnak a betegek mert számítanak rá hogy meggyógyulhatnak így felkelhetnek az ágyból [MASK] az almát és a körtét. *The doctor whom the patients wait for because they hope they can be cured thus they can get up from bed [MASK] apples and pears.*	S/3	S/3	2/5
10.	Az orvos akit várnak a betegek mert számítanak rá hogy meggyógyulhatnak és felkelhetnek az ágyból mivel lesz már annyi erejük [MASK] az almát és a körtét. *The doctor whom the patients wait for because they hope they can be cured thus they can get up from bed since they have enough energy [MASK] apples and pears.*	S/3	infinitive	0/5

Table I. Results of the first sentence in each subset

Typically, there was great variation in the results within one subset, which tended to follow a pattern depending on whether the main subject was singular or plural. There was, however, no variation in the results of the first subset. For all but one sentences, the BERT model predicted a period instead of the expected verb meaning that it evaluated sentences like “The doctor whom the patients are waiting for.” as whole. Interestingly, it chose the period as the most likely token despite its presence in the original input, meaning that it did not pay much attention to the right side of the masked word.

Its accuracy however improved significantly when an adverb was added in front of the masked word. It tended to be slightly more accurate with the adverb “here” as opposed to the adverb “very”, suggesting that spatial dimension is treated as potentially more independent from its direct context as the intensifier adverb. The model guessed accurately for all the singular subject sentences; however, it missed both plural object sentences in both subsets. The correct prediction still appeared among the top five suggestions for the plural subject sentences with the “here” adverb. This categorical difference between the singular and the plural subject sentences exhibited a quite consistent pattern throughout the testing.

In the next subset I was looking to find out whether the accuracy of the model improves in the case of the intensifier adverb if I add a direct object clause after the masked word. Interestingly, the overall number of the correct top guesses decreased, only two sentences were completed grammatically correctly. This time the model made a mistake even in the case of a singular subject sentence and in one plural subject sentence, it could not even guess the class of the masked word, mistaking it for an adjective.

To continue to explore the effect of a direct object clause on the model’s accuracy, I removed the “bad influence” adverb (note that in Hungarian “really” and “very” are one word) and added an embedded clause in front of the masked word instead, while still preserving the direct object clause linked to it. The embedded clause was a verb phrase referring to the subject of the subordinate relative clause. As I was expecting, the accuracy of the model decreased even more, meaning that a verb phrase as an embedded clause before the mask is even more of a challenge to it in terms of treating the two constituents separately. In the case of the plural subject sentences, the model was unable to tell the class of the missing word and opted for a conjunction and a preposition, even though these choices rendered both sentences ungrammatical.

Since one obvious possible explanation for this decrease was the fact that these sentences were the longest by far, in the next subset I was also looking at the effect of sentence length on accuracy. This time I included the adverb “really” as well as the embedded clauses of the previous subset, along with the direct object clauses. Contrary to my expectations, the model’s accuracy was visibly improving. Besides the overall accuracy, it has even managed to guess the form of the masked verb of a plural subject sentence correctly for the first time during this experiment. This has led me to believe that sentence length itself, at least within the window of these six subsets was not strongly influential in hindering or helping the model find the correct syntax of the sentence.

In the fifth subset it was suggested that the model's accuracy is decreasing in the presence of an embedded clause that refers to the subordinate relative clause of the main subject. For the seventh subset I was looking to find whether it is the mere presence of such a clause or the subject it refers to influences the model's syntactic understanding. For this subset, I have changed the dependency of this clause, linking it to the main object of the sentence. The accuracy of the model improved significantly, only one of the sentences was not guessed correctly, producing the best overall result so far. Nevertheless, it is not clear whether the result improved because the overall understanding of the model improved or just the proximity of an embedded clause with the same form as the masked verb gave an influence in the correct way.

To further explore this question, I designed the next subset of sentences by inserting the adverb "really" in between the embedded clause and the masked verb. The accuracy of the guesses slightly decreased. In one sentence, the model mistook the masked word for an adjective, revealing that it only took the direct environment of the mask as context.

Finally, in the last two subsets, I was purely looking at distance as a factor. In these subsets I wanted to see the distance the model can "tolerate" between a verb and its subject. The ninth subset contained sentences with four recursively embedded clauses between them, whereas the tenth subset contained five. All of them were dependents of the subject of the relative clause, not the main subject. Interestingly, for two out of the three singular subject sentences, the unnaturally long distances and stacking of subclauses posed no problem, the model was still able to correctly guess the missing word. This suggests that there might be some understanding of recursivity within the model, and when this feature is exaggerated, for example with five subclauses, it is able to "regain" its focus. The model seemed to struggle however with the sentences that had plural main subjects, not even being able to identify the class of the missing element.

Conclusion

Although the experiment was small in scale, there were definitely patterns emerging that suggested how the HUBERT model's syntactic-grammatical understanding changes depending on the sentence structure. In this section I would like to list the key observations I have made during the experiment and some potential explanations.

First, across all subsets, the model seemed to be unable to guess the correct structure if the main subject and hence the expected correct verb form was plural. Since the focus of the experiment was the presence of an attractor noun, I could not identify a possible reason, why the switch to plural posed challenges for the model in the test set. This pattern alone would be worthy for further research in a separate experiment.

Second, the model's accuracy significantly declined when dealing with a main subject that is inanimate. Even though it was a singular subject thus it would have been easier for it

to understand, the model visibly struggled predicting the correct verb for the subject “car” as opposed to the subjects “doctor” and “mother”. Although more extensive experimentation would be required to come to any conclusion about this phenomenon, one potential explanation might be that the model tends to pay more attention to animate nouns over the inanimate ones when deciding on the correct form of a verb.

The overall least accurate predictions occurred in the first subset, where zero guesses were correct. This showed that the model did not pay sufficient attention to the period at the end of the sentence, since its top suggestions were all but one sentence-ending symbols despite there already being one. This subset was also the one that gave the most “creative room” for the model, since the mask covered a standalone verb phrase, without any adverbs or clauses attached to it, that would narrow its potential meaning. This suggests that the model does not do well in sentences where it has to be too creative and would rather resort to well defined closed classes such as punctuation marks.

The most accurate top guesses occurred in the subset 7, where the masked verb was preceded by a relative clause of a verb phrase referring back to the main subject and proceeded by a direct object clause. The presence of a direct object clause has proven to be an enhancing factor across all the test set, aligning with my previous suggestion, that the model performs better with less room for creativity. The accuracy of the predictions in this subset also suggests that sheer proximity can play a role in the model’s understanding, since these sentences included a verb phrase with the correct verb form, even though syntactically, these verb phrases were not dependent on the masked verb.

Finally, I have mentioned that Hungarian language tends to concatenate the verb “are” and the adjective it relates to (e.g. instead of “they are elegant”, we simply say “they elegants”). The model did not seem to struggle with this concatenation, the accuracy remained independent of whether the subclauses contained implicit verbs or explicit verbs.

In conclusion, there were observable patterns in the model’s ability to predict the correct verb form of the masked verbs despite the presence of attractors. It is important to reiterate however that the scale allowed for more of a first impression and discovery of potential areas worthy of further research rather than finite conclusions.