

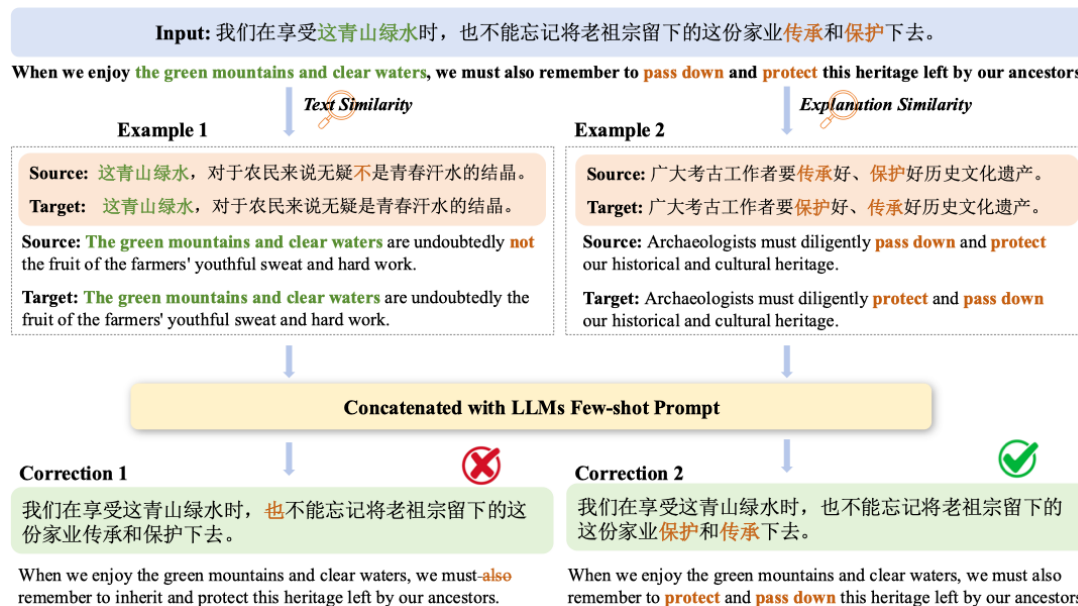
RE²: Improving Chinese Grammatical Error Correction via Retrieving Appropriate Examples with Explanation

**Baoxin WANG, Yumeng LUO, Yixuan WANG
Dayong WU, Wanxiang CHE, Shijin WANG**

Frontiers of Computer Science, DOI: [10.1007/s11704-025-41399-w](https://doi.org/10.1007/s11704-025-41399-w)

Problems & Ideas

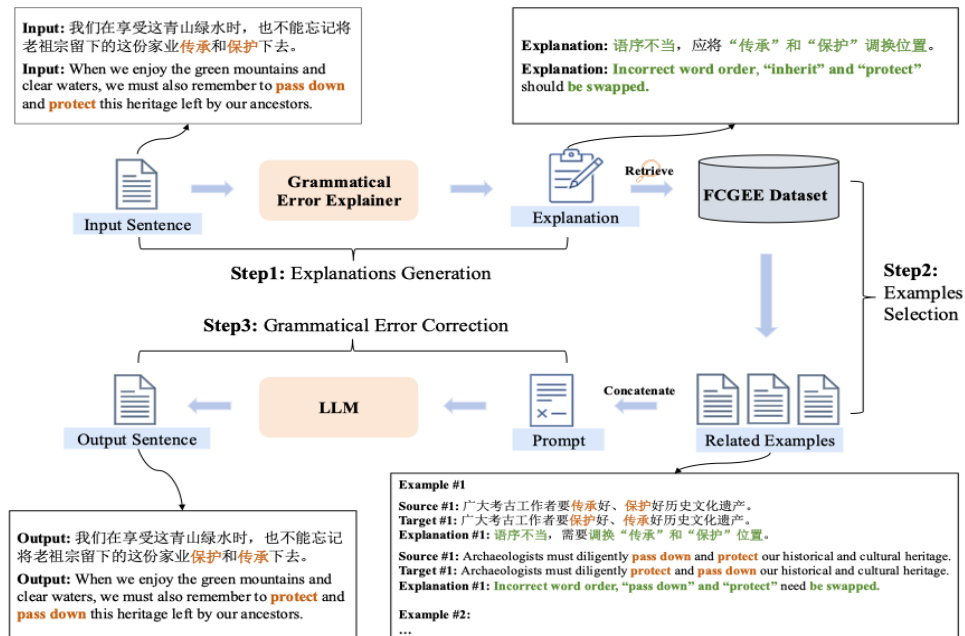
- Problems of Examples Retrieval Method for GEC:
 - Sentences with similar surface structures may involve fundamentally different types of grammatical errors, leading to mismatched or irrelevant reference examples.
 - Text-similarity-based retrieval fails to account for the underlying error patterns or linguistic rules, limiting the model’s ability to generalize error-specific correction strategies.
- Ideas: Retrieve similar examples based on the explanation of a sentence’s grammatical errors instead of directly using the input sentence itself.



Examples of retrieving reference examples. Example 1 is retrieved using text similarity of the input sentence, while Example 2 is retrieved using the grammatical error explanation of the input sentence. The green-colored text is the overlapping part between the input and the text of Example 1, while the orange-colored text represents the corresponding errors and correction results. The overlapping part found between the reference example and the input using the text similarity method is “这青山绿水” (the green mountain and clear waters), while the explanation similarity method can identify examples with the same errors in the order of “传承” (pass down) and “保护” (protect)

Main Contributions

- Contributions:
 - Construct a high-quality dataset of Chinese grammatical error explanations;
 - Propose the RE² method and explore the use of grammatical error explanations for retrieving appropriate reference examples;
 - Our method achieves the best performance on two datasets, and we conduct an in-depth analysis of grammatical error explanations.



The green-colored text represents the overlapping part of the explanation texts, while the orange-colored text corresponds to the error and correction parts in the example.