Title of The Inventions: Text Segmentation Method, Text Segmentation Apparatus, Electronic Device, and Computer-Readable Storage Medium

Abstract

The present disclosure relates to computer technology and provides a text segmentation method, a text segmentation apparatus, an electronic device, and a computer-readable storage medium. The method comprises: acquiring a plurality of sentences obtained after sentence-level segmentation of a target text; sliding a preset sliding window multiple times by using the first sentence of the plurality of sentences as a starting point and using the last sentence of the plurality of sentences as an end point, and semantically segmenting a plurality of target sentences located in the preset sliding window after each sliding operation to obtain at least one semantic segment; and clustering, for semantic segments obtained by each segmentation, the semantic segments obtained by the current segmentation and semantic segments obtained by a previous segmentation to obtain a target semantic segment to which each of the target sentences in the preset sliding window after the previous sliding operation and after the current sliding operation belongs, until a semantic segment to which each sentence of the plurality of sentences belongs is obtained. The present disclosure can ensure semantic coherence and improve the accuracy and reliability of question answering in a RAG system.

FIG. 2