Abstract

The invention provides a computer-implemented method for analyzing a piece of text comprising Chinese characters. The method comprises the steps of truncating the piece of text into a plurality of first block units each having a first predefined number of N characters, where N is an integer and is greater than or equal to one; determining, for a selected character from the N characters of each of the first block units, one or more radicals forming said selected character; identifying, from the one or more radicals forming said selected character, one or more semantic radicals by comparing the one or more radicals with a database comprising semantic radicals and their associated meanings, and determining one or more meanings of the one or more semantic radicals in relation to said selected character; categorizing the plurality of first block units into one or more category groups based on the determined one or more meanings of the one or more semantic radicals of the selected character of each of the first block units; and computing a number of the first block units categorized in the respective one or more category groups indicative of one or more characteristics of the text.

FIG 1