ABSTRACT OF THE DISCLOSURE

CHINESE GENERATION APPARATUS FOR MACHINE TRANSLATION

The Chinese generation apparatus for machine translation utilizes statistic data for Chinese generation instead of a large number of semantic, grammatical related rules, the statistic data being retrieved from a tagged Chinese corpus. In order to overcome the problem that a large number of semantic or grammatical rules are needed in Chinese generation, the constitution of a Chinese sentence is divided into basic sentence pattern item and other items. The probability of the possibly generated basic sentence pattern of each Chinese verb and the location probability of the other item in the basic sentence pattern are initially retrieved from the Chinese corpus and are stored in the buffer unit. The best basic sentence pattern and the location of the other items can be found in accordance with each constitution item of the intermediate form of Chinese and the statistic information in the buffer unit with the use of branch-and-bound method, so that the most suitable Chinese sentence can be generated.

Fig. 1