Title: GENERATING CHINESE LANGUAGE COUPLETs

Abstract: An approach of constructing Chinese language couplets, in particular, a second scroll sentence given a first scroll sentence is presented. The approach includes constructing a language model, a word translation-like model, and word association information such as mutual information values that can be used later in generating second scroll sentences of Chinese couplets. A Hidden Markov Model (HMM) is used to generate candidates. A Maximum Entropy (ME) model can then be used to re-rank the candidates to generate one or more reasonable second scroll sentences give a first scroll sentence.
A CLASSIFICATION OF SUBJECT MATTER
IPC(8): G06F 17/00 (2007.01)
USPC: 706/45
According to International Patent Classification (IPC) or to both national classification and IPC

B FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
USPC 706/45

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPG 706/45, 50, 58, 715/523, 536, 703, 700/47, 48, 49

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Electronic data base PubWEST (PGPB.USST.USOC.EPAB.JPAB), Google Scholar
Search Terms Used couplet and natural language and automatic translation and unigram, digram or ttriga, Chinese couplet or Chinese parser or automatic Chinese translation etc

C DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category*</th>
<th>Citation of document with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>US 2003/0083861 A1 (WEISE D N ) 01 May 2003 (01 05 2003) (para [0056],[0057],[0058],[0059])</td>
<td>3-4, 5</td>
</tr>
<tr>
<td>Y</td>
<td>US 5,900,746 A (TING H A ) 27 July 1999 (27 07 1999) (Fig 3, 4, col 7, Ins 52-55, col 7, In 65 to Col 8, In 13, col 9, In 65 to col 10, In 10)</td>
<td>5, 9-11</td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of Box C

* Special categories of cited documents
'A' document defining the general state of the art which is not considered to be of particular relevance
'E' earlier application or patent but published on or after the international filing date
'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
'O' document referring to an oral disclosure, use, exhibition or other means
'P' document published prior to the international filing date but later than the priority date claimed
'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
'X' document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
'Y' document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
'Z' document member of the same patent family

Date of the actual completion of the international search
23 February 2007 (23 02 2007)

Date of mailing of the international search report
07 MAY 2007

Name and mailing address of the ISA/US
Mail Stop PCT, Att x ISA/US, Commissioner for Patents
P.O Box 1450, Alexandria, Virginia 22313-1450
Facsimile No 571-273-3201

Form PCT/ISA/210 (second sheet) (April 2005)

Authorized officer
Lee W Young
PCTHelpdesk 571-272 4300
PCTOBP 571-272-7774