



圖三 中研院詞知識庫小組的分詞和詞性標記程式

<http://ckipsvr.iis.sinica.edu.tw/>

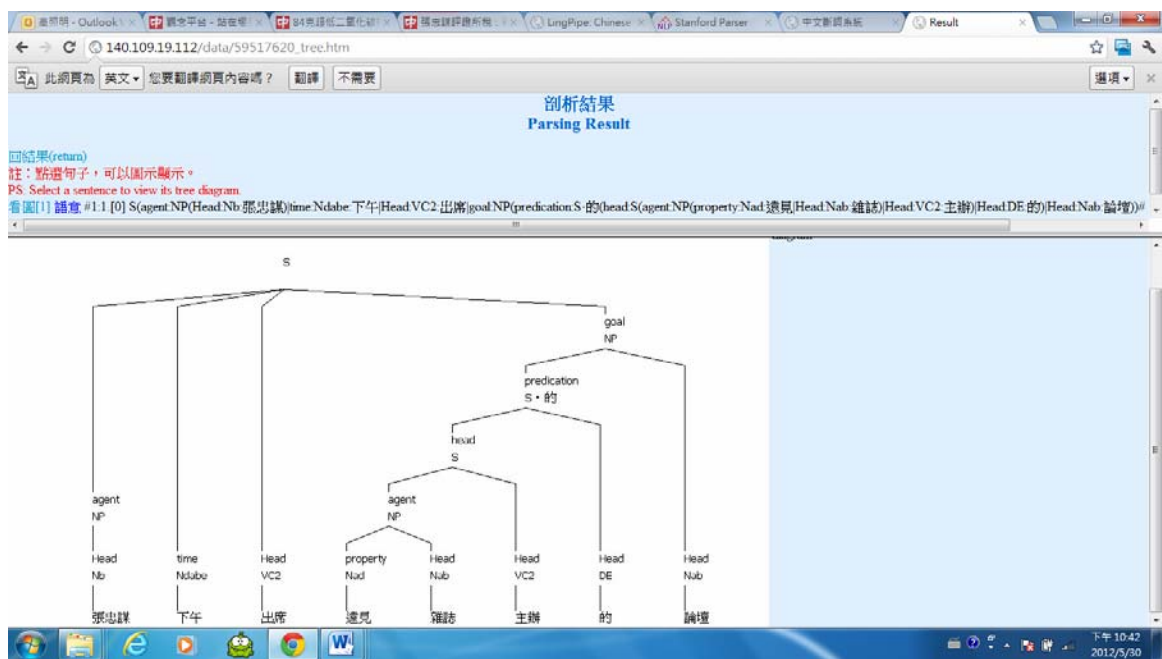
六、語法剖析器(parser)：程式自動將輸入的句子的句法層次結構標示出來。語法剖析器可以分成兩種，完全剖析和部分剖析(partial parse)。近年來興起能判斷依存關係的語法剖析器,如英文的 Minipar 及 Stanford Parser。瑞典 Lund 大學以 Mate-tool 為基礎發展簡體中文的語法剖析器提供程式碼供研究人員下載。

Sinica Treebank 與 Penn Chinese Treebank 最大的差別在於結構樹的語法單位不同。前者以標點符號作為分隔不同結構樹的單位，因此一個結構樹很多時候只是一個詞組（如 PP, NP）而不是一個完整的句子。而後者除小部分結構樹是句子的片段（以 FRAG 標示）大部分的結構樹是完整的句子(sentence)(以 IP 標示)。另外 Sinica Treebank 語法結構採取中心語主導原則（Head-Driven Principle），註明中心語(Head)和其他成分（如附加語）的語法和語意訊息，表達出句子中詞和詞之間的語法結構和語意角色關係，而 Penn Chinese Treebank 並沒有中心語與語意角色的訊息，而是在詞組上加註如主詞 SBJ 受詞 OBJ 等語法功能的方式來取代。

如（圖五）所示，中研院的中文句法樹庫的 terminal node 是詞，詞上方有詞性標記和中心語（head）這類的語法訊息，構成詞組的結點(node)有詞組標記和語意角色等語意訊息。



圖四 中研院詞詞知識庫小組中文剖析器的輸入介面



圖五 中研院詞知識庫小組中文剖析器的輸出介面

圖六是 Stanford Parser 中文剖析器的輸入介面。圖七 Stanford Parser 中文剖析器的輸出最重要的部分是句子的語法結構樹和語法依存關係。語法結構樹顯示詞組之間的語法關係如 NP 是名詞組，VP 是動詞組，IP 是句子。圖七最下方顯示詞與詞與詞之間的語法依存關係。



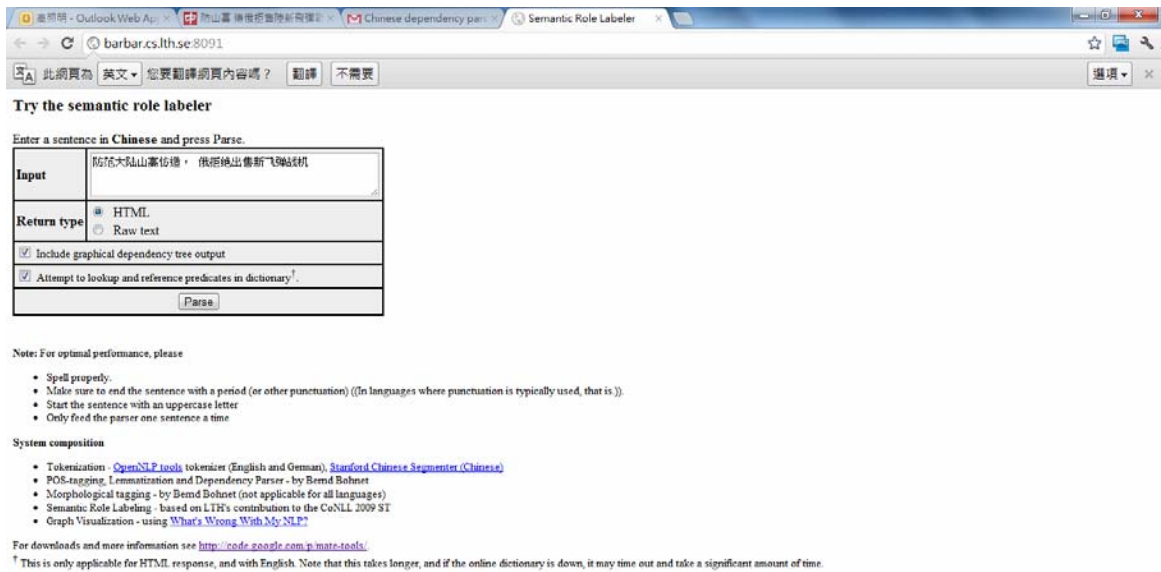
圖六 Stanford Parser 中文剖析器的輸入介面

<http://nlp.stanford.edu:8080/parser/>

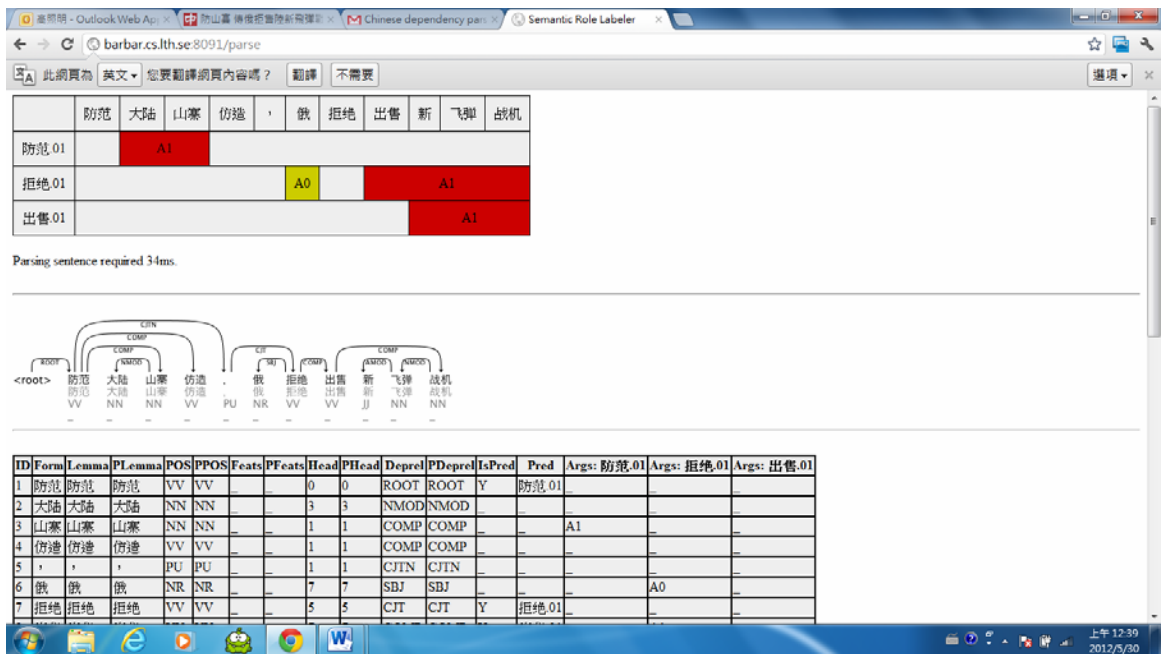


圖七 Stanford Parser 中文剖析器的輸出介面

<http://nlp.stanford.edu:8080/parser/>



圖八 瑞典 Lund 大學以 Mate-tool 為基礎發展簡體中文的語法剖析器的輸入介面
<http://barbar.cs.lth.se:8091/>



圖九 瑞典 Lund 大學以 Mate-tool 為基礎發展簡體中文的語法剖析器的輸出介面
<http://barbar.cs.lth.se:8091/>

七、線上機讀辭典：教育部國語會出版的國語辭典，除了解釋，並有例句，相似詞，相反詞。