二、研究計畫內容

(一) 摘要

本研究欲透過「自動問句生成閱讀理解選項」為應用目標，欲透過提

升問句生成過程中關鍵字篩選之難度及複雜度，來增進自動問句生成閱讀理解模型對未來教育環境的幫助。

(二) 研究動機與研究問題

近年來，因為自然語言處理領域的技術較成熟，使得自動問題生成成為開發未來學習系統理想的工具。如今，自動問句生成之閱讀理解模型已廣泛應於多種教育用途，像是學術寫作支持(academic writing support)[1]、閱讀理解評估(reading comprehension assessment)[2]、教育聊天機器人(educational chatbots)[3]。但上述的系統中，是利用文章中部分段落來產生文法通順且合理的問題來達到教育目的，然而，這些系統卻只依賴簡單的挑選方式來選擇關鍵字或片語作為問題的答案。也因為作為答案的關鍵字挑選方式較為簡單，造成在問句生成時，無法真正透過這些問題來評估學生在學習上的效果，這也是現在自動問題生成系統還無法取代人工出題的其中一項原因。

因此，關鍵字的選擇是一項讓自動問句生成擴展到具有預先指定答案的數據集之外，並且能應用於現實生活中的重要步驟。

現今，市面上較為廣為人知且強大的ner (Name Entity Recogonition) API 有NLTK、spaCy，但他們就僅能針對文章中的人名、地點、時間…等，進行標記，但我們知道如果要完成一個完整的自動出題系統，並能確保學生真的完整理解一篇文章的內容，我們不應該將考題侷限於人名、地點…等較短專有名詞上。因此，本專案想開發一套能篩選出閱讀文本中一段文字做為答案的系統，如此不僅能增進自動問句生成系統的問題難度，更能深入了解學生是否真的完全了解此篇文章中的內容。

以圖一為例，Text 1 的內容為美國布希總統在選舉時，請到了美國著名的民調公司Gallup 來為自己做民調，而民調的時間是從1992年11月至1993年1月之間製作的，但我們使用spaCy這款套件來做ner，我們發現雖然spaCy為我們分別挑出了「November 1992」和「January 1993」兩個關鍵字，也就是spaCy認為這兩個關鍵字為獨立的個體，但以語意上來說，「November 1992 through January 1993」為一區間，如果將兩個時間點分割成兩個獨立的個體，對於讓學生了解文章的內容，沒有幫助，因此本專案將著手解決市面上這些套件所造成的問題。

圖一 :

Text 1: George H.W. Bush averaged an 84% approval rating among Republicans in an average of November 1992 through January 1993 Gallup

polls.

Result : November 1992 (DATE)

January 1993 (DATE)

(三) 文獻回顧與探討

(四) 研究方法與步驟

(五) 預期結果

(六) 參考文獻

[1] Ming Liu, Rafael A Calvo, and Vasile Rus. 2012. G-Asks: An intelligent automatic question generation system for academic writing support.

[2] Jack Mostow and Hyeju Jang. 2012. Generating diagnostic multiple choice comprehension cloze questions. In Proceedings of the Seventh Workshop on Building Educational Applications Using NLP. Association for Computational Linguistics, 136–146.

[3] Nasrin Mostafazadeh, Ishan Misra, Jacob Devlin, Margaret Mitchell, Xiaodong He, and Lucy Vanderwende. 2016. Generating Natural Questions About an Image. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 1802–1813.

DOI: http://dx.doi.org/10.18653/v1/P16-1170

(七) 需要指導教授指導內容