Universal Transformer

1. 我要介紹的是Universal Transformer，一般的Transformer雖然在翻譯的方面表現的很好，但在演算法測試時卻不太好，因此Universal Transformer是一個可以在翻譯和演算法測試都很好的Transformer模型。另外Universal Transformer在每層的參數都是一樣的如圖1所示。

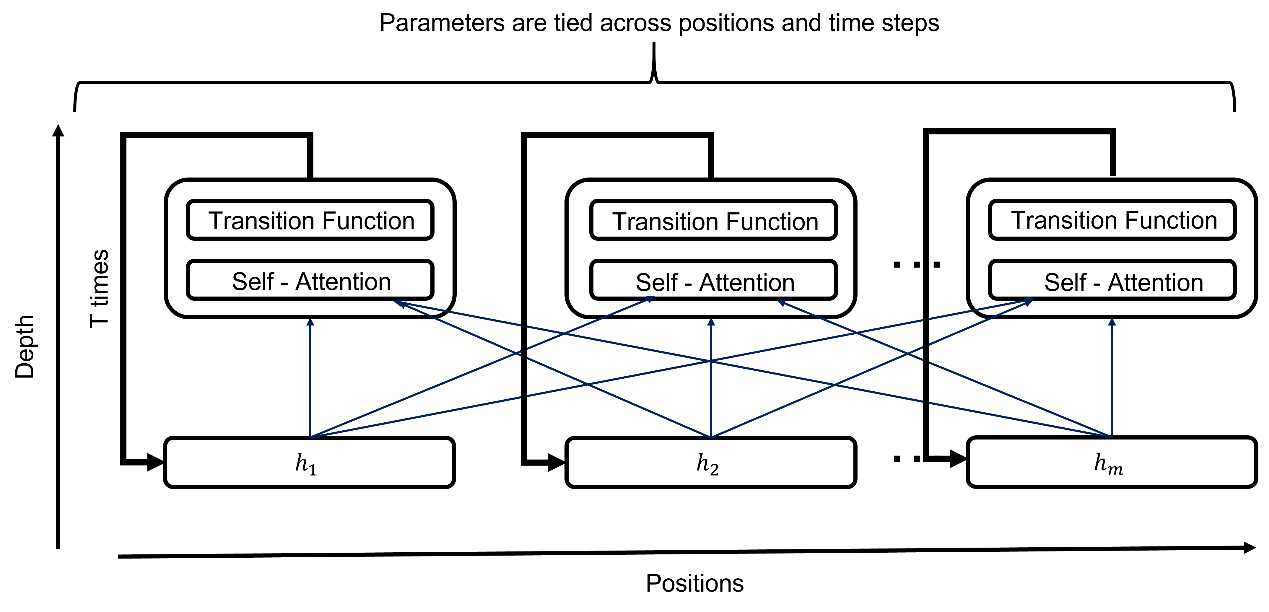


圖1. Universal Transformer 架構圖

1. Universal Transformer的優點我認為在於:

* 因每層的參數都一樣，因此其比傳統的Transformer在計算上更有效率。
* 其增加了Dynamic Halting的機制，這意味著該模型可以根據每一步的輸入和輸出調整其計算，使其能夠處理各種序列長度和依賴性，是模型的表現上升。

參考資料 :

1. [(874) [TA 補充課] Transformer and its variant (由助教紀伯翰同學講授) - YouTube](https://www.youtube.com/watch?v=lluMBz5AoOg)
2. https://ai.googleblog.com/2018/08/moving-beyond-translation-with.html