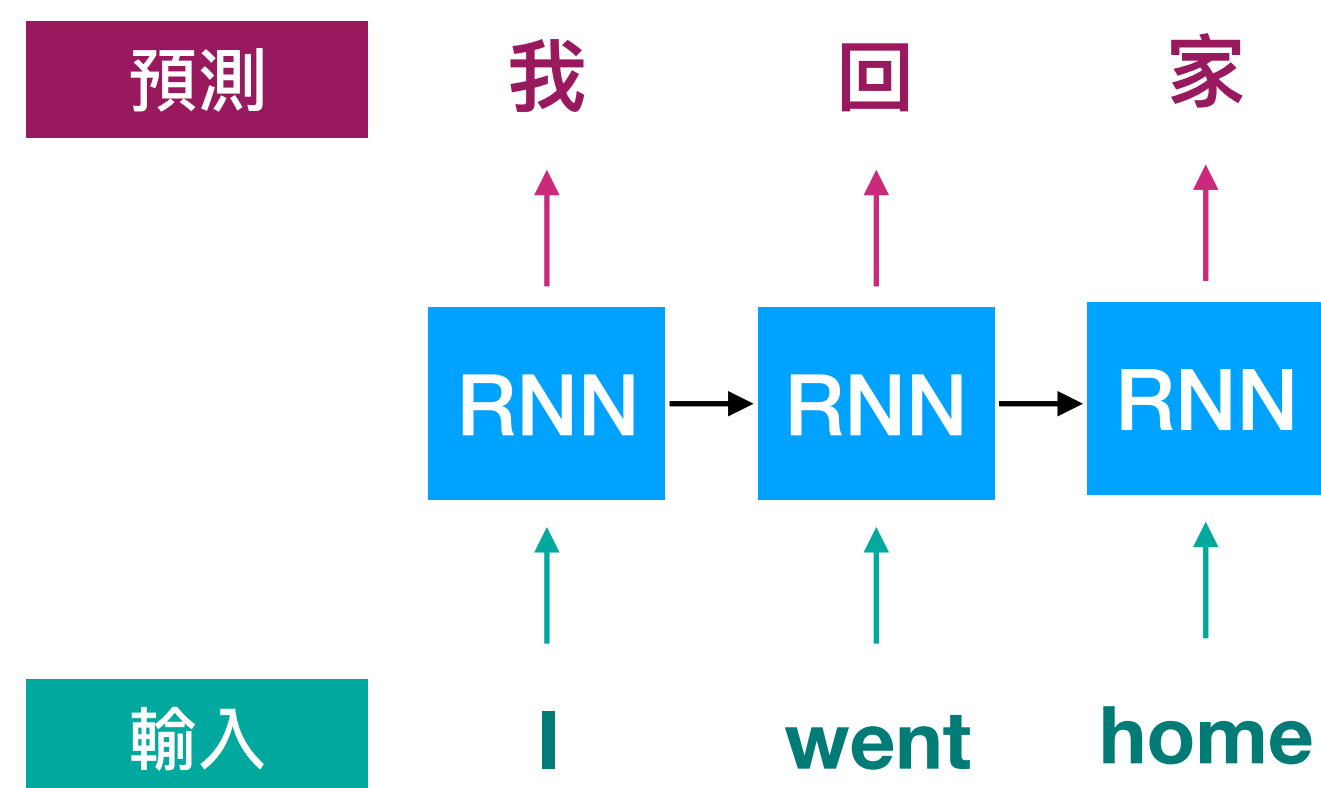
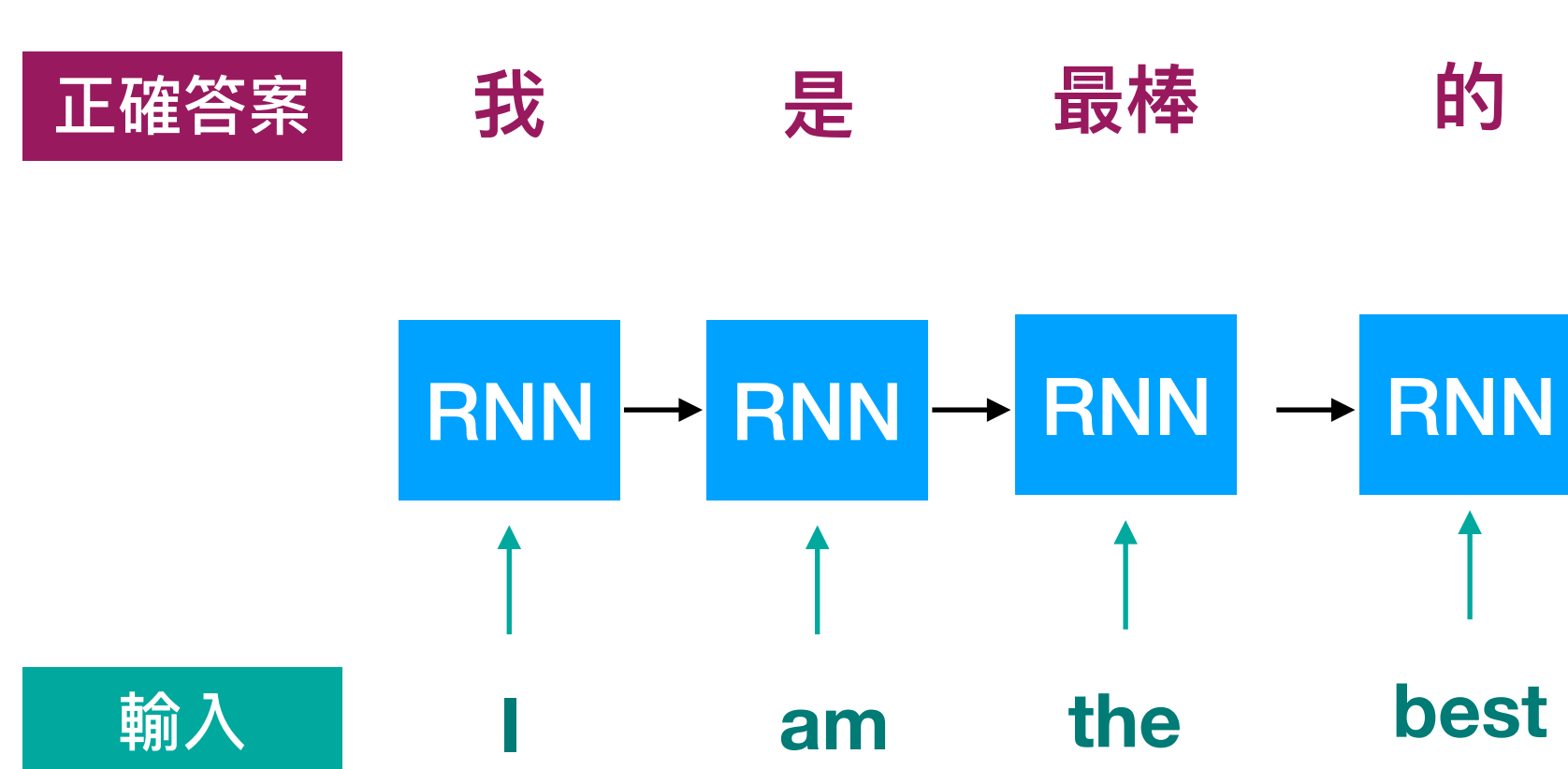


機器翻譯問題

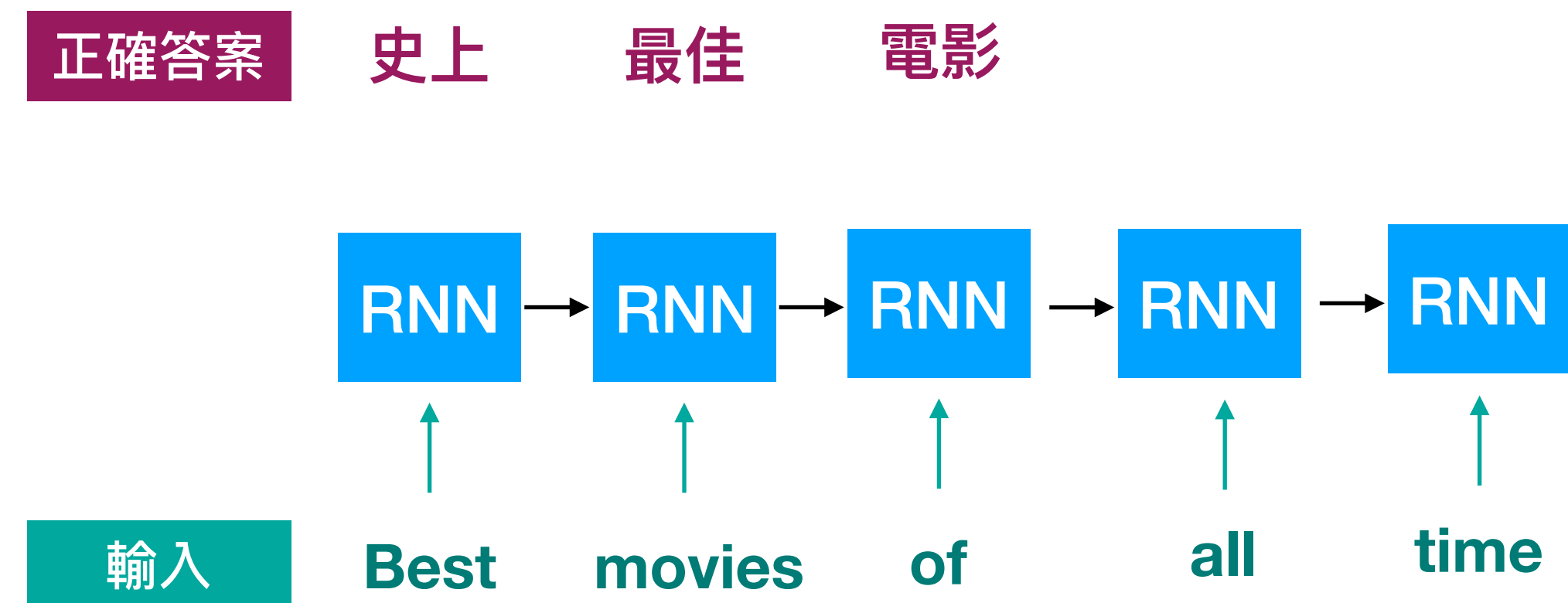
既然我們已經學會 RNN 了，我們可不可以讓 AI 學會翻譯呢？假設以我們的 RNN 來做翻譯有可能嗎？



問題



輸入和輸出對不起來！因為the根本沒意義，應該是看完 best 再做翻譯



一樣對不起來！因為中文只有三個，但英文有五個，而且剛好英文和中文是反過來的

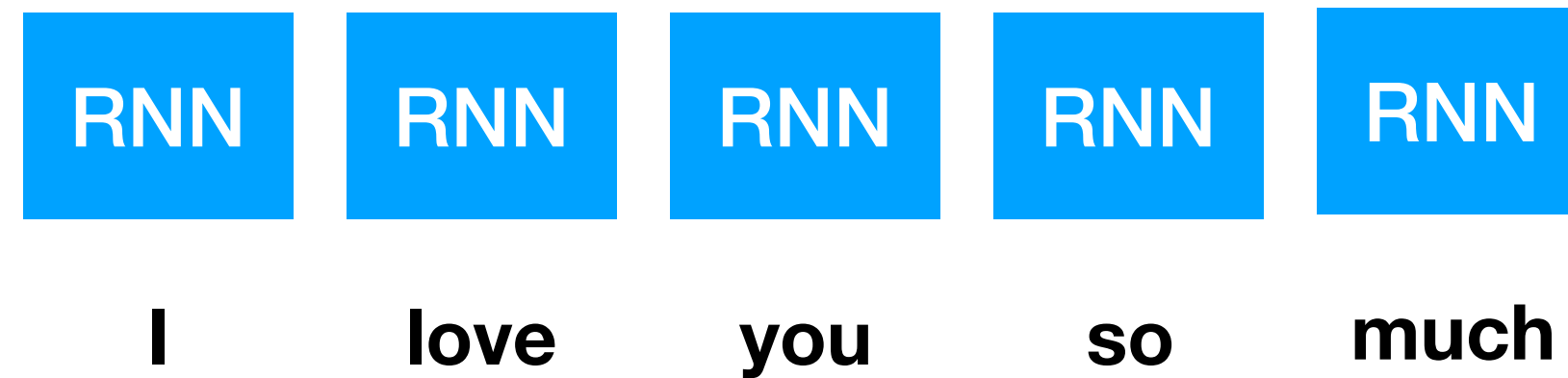
解決辦法

Seq2Seq

就四個字，『看完再說』不像之前每看一個字或詞就開始產生，我們完整看完整個文章，再依照剛剛看完的記憶產生翻譯

Encoder (編碼器)

將一個複雜的文章壓縮成一個語意



Decoder (解碼器)

將語意解壓縮成複雜的文章



訓練模型

建議方式

Decoder (解碼器) with Teacher Forcing

語意



Teacher Forcing :

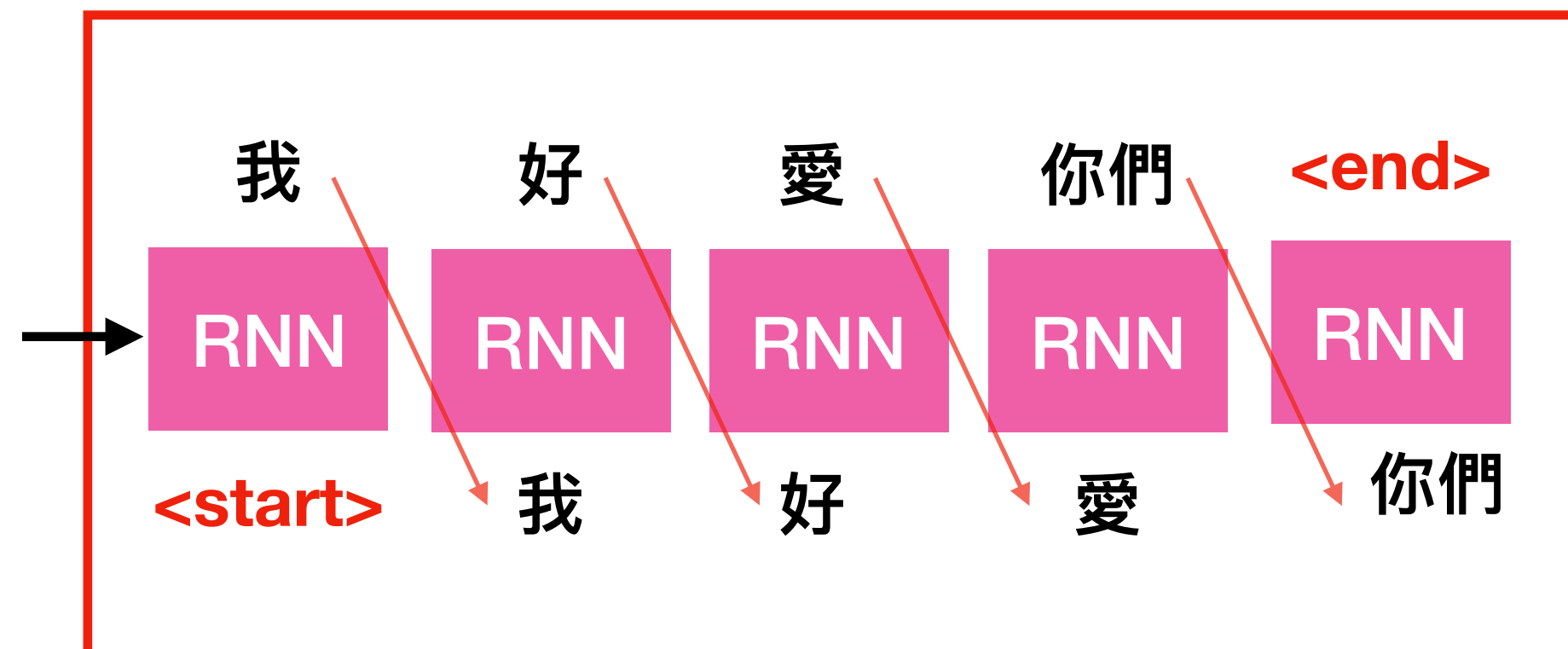
第一種Decoder訓練方式就如同我們之前的Char RNN，輸入是第 1~ (n-1) 個字，輸出是第 2 ~ n 個字，優勢是就算上一個字預測錯，我仍然可以透過正確的輸入來調整記憶並且預測下一個字

這裡要注意，因為第一個字也要預測，所以我們第一位位置帶入一個你自訂的<start>符號代表文章的開始，預測第一個字，當然也要有結束，所以當預測出<end>自訂符號的時候結束句子

較不建議

Decoder (解碼器) without Teacher Forcing

語意



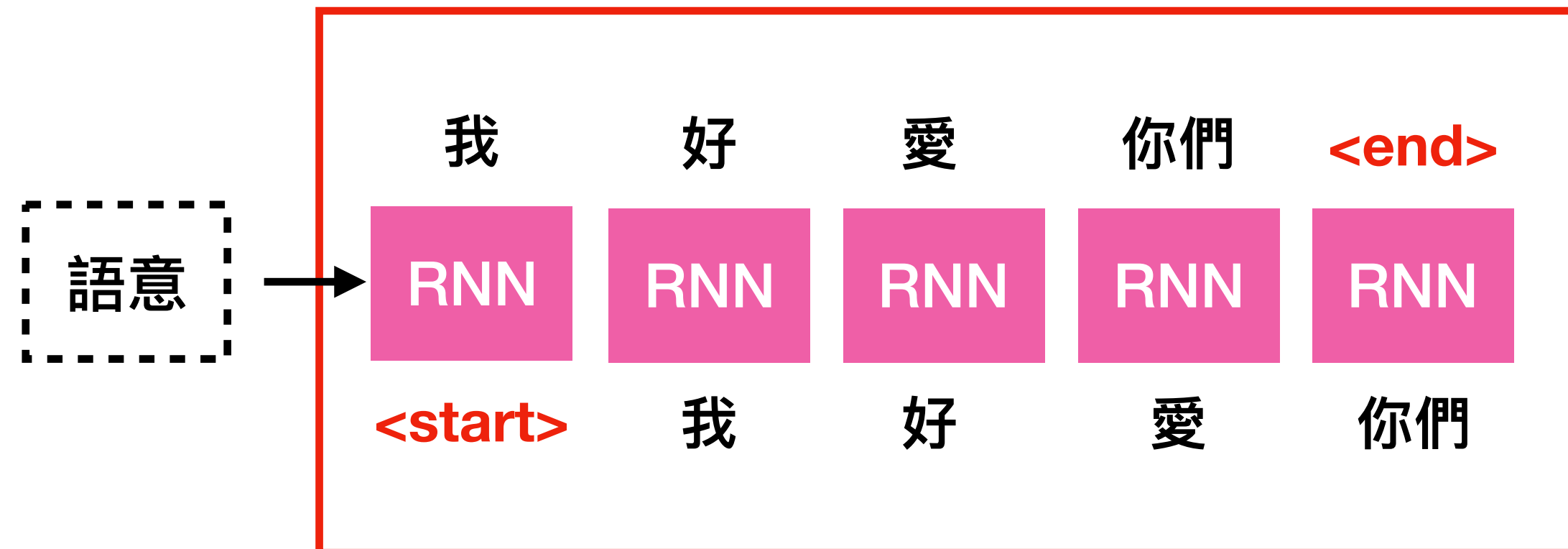
無Teacher Forcing :

直接把輸出待到下一個輸入，比較建議是因為沒有 Teacher Forcing 的話，你可能會因為第一個預測錯，後面就蝴蝶效應，越錯越多，比較不建議

訓練模型

建議方式

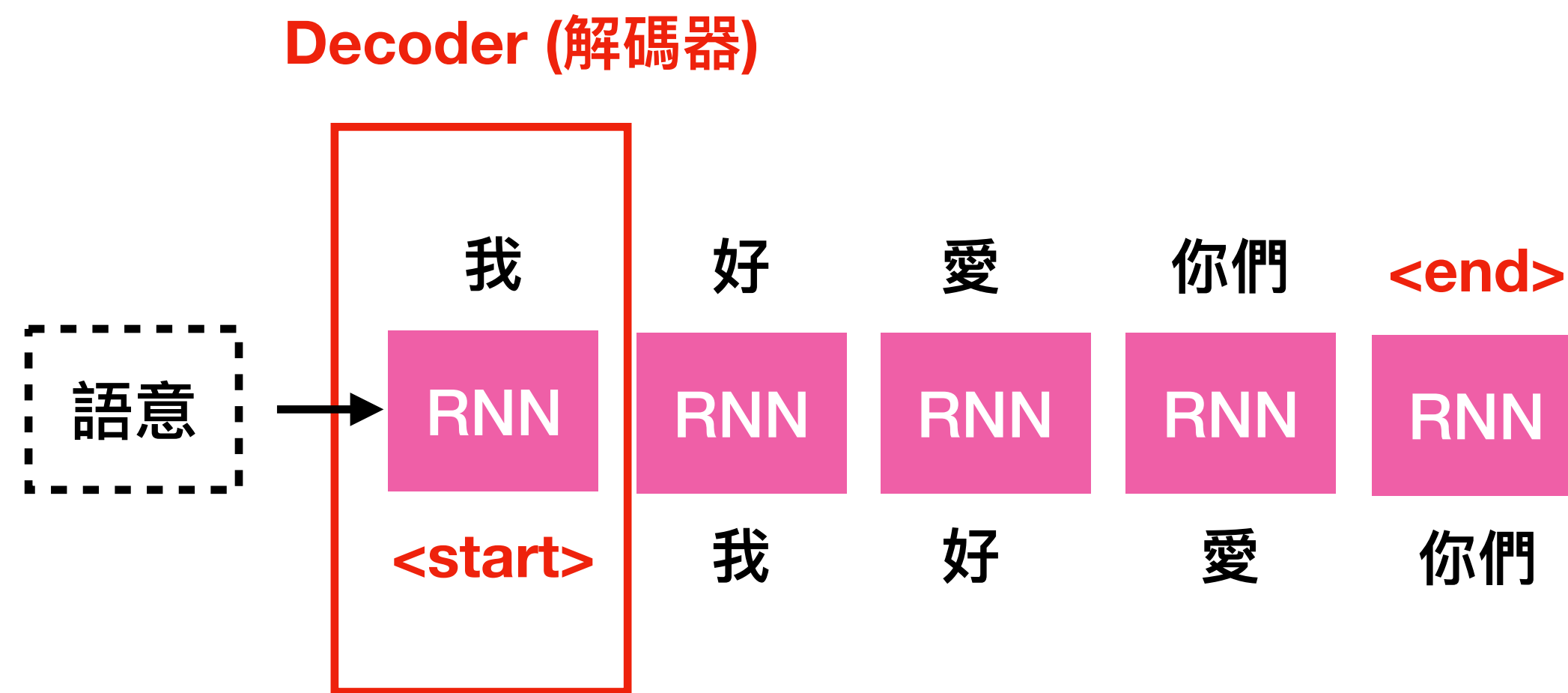
Decoder (解碼器) with Teacher Forcing



Teacher Forcing :

第一種Decoder訓練方式就如同我們之前的Char RNN，輸入是第 1~ (n-1) 個字，輸出是第 2 ~ n 個字，優勢是就算上一個字預測錯，我仍然可以透過正確的輸入來調整記憶並且預測下一個字

推論模型



推論的時候其實就是 Char RNN，不斷輪迴預測下一個詞，只帶入 <start> token，無窮迴圈不斷預測下一個詞直到看到 <end> token 為止