

一、 因為最近剛好看了一部電影《降臨》。

這讓我聯想到 AI 在未來 20 年，有可能具備「自動推演新文字或語言系統並能對應已有語言結構」的能力。現階段的模型僅能在已知語料中進行統計比對，無法理解語言背後的邏輯與語意。然而，若結合符號學推理、跨越不同邏輯之間的學習與世界模型，AI 將能在完全未知的語言資料中自動推測語法規則、語意結構與概念映射，進而建立新舊語言間的對應關係。這項能力對人類文明意義重大，不僅能破譯失傳文字、理解外星訊息，亦可協助語言復原與文化保存。它象徵 AI 由「模仿語言」進化為「理解符號」，使知識傳遞不再受限於語言邊界。

二、 此能力的實現將主要依賴非監督式學習與強化學習的結合。

首先，AI 必須在缺乏標註資料的情況下，自主從未知文字樣本中學習符號之間的統計關係與結構規律，因此核心為非監督式學習。這使模型能歸納語法模式、語素組合與可能的語意分佈，類似人類語言學家在沒有字典時推測語義的過程。資料來源為新語言的符號序列、文字圖像或語音樣本，而目標訊號則是符號間的內在一致性與結構可預測性。

其次，若要進一步將新語言對應到舊語言系統，AI 需透過強化學習在假設-驗證過程中不斷修正映射。環境回饋可由語義對齊準確度或可解釋性評估產生，使模型逐步逼近合理的翻譯與概念對應。兩者結合後，AI 才能不依賴人工標註，自主探索符號與意義間的深層關聯。

三、 模型化構想

作為第一步的簡化問題，可以設計一個任務：讓 AI 學會在沒有任何語法標註的情況下，從大量被隨機打亂語序的句子中，推斷出原始語言的基本文法結構。例如輸入「飯吃我喜歡」「戰爭打我討厭」等句子(內部的文字不會是我們已知的文字，而是透過某種簡單的規律去生成，但都不是已知單字)，AI 需自行學習出「主詞—動詞—受詞」等語法規則，並重構出合理語序。此問題在概念上模擬了 AI 對未知語言結構的歸納過程。

可測試性在於模型能否正確還原語序，或能否在新句中預測出合理的語法關聯。

數學與技術上可運用機率圖模型 (Probabilistic Graphical Models)、自注意力結構 (Transformers) 與非監督語言建模 (Unsupervised Grammar Induction)，以統計關聯與語意相似度為基礎推論出語法規則。