**機器翻譯:**

詞彙面向:一個字詞從最一開始機器翻譯以來慢慢地演化答案越來越符合句意

結構面向:有點像是句子重組，例如英翻中的過程當中每次不斷地進步調換介系詞片語的位置，得出一個最流暢的句子

**問答系統:**

挑戰:分析問題、內容，計算支持、反駁資訊的信心度，因此牽涉大範圍的技術

已有應用:Siri、智慧音箱

**意見探勘:**

文本分析🡪對不同資料進行文件探勘

**典型問題**:

歧義性(詞彙層次、語法層次、語義層次)、容錯(打錯字、語法錯誤、尚未收錄的未知詞)、強健性(領域改變、網路語言、表情符號)

---------------------------------------------談自然語言處理-------------------------------------------

**解決上述問題🡪掌握共通性:**

詞性類別(名詞、動詞、形容詞、副詞…)、語義類別(同義詞)、句法類別(分析句子的主詞、名詞片語、介係詞片語…)、相依類別(詞跟詞之間的相依關係)、

言談類別(句子與句子之間的關係 ex.時序、因果、轉折、推展)、意見類別(正、負面)、情感類別(像是打字聊天時會打表情符號、句子中帶有情感、情緒)、立場類別(核能發電、廢除死刑、同性婚姻…)、電腦也有智慧，他必須要把一些有規律的現象記錄下來，那麼**分類**就是一種很重要的機制

以上詞彙和類別都以符號呈現、匹配是基本符號計算

------------------------------------向量---------------------------------------

分布式表示🡪語境(上下文)🡪上下文範圍界定🡪共同出現🡪透過向量計算夾角計算出兩個詞彙的關係🡪進一步可利用cos值判定兩個詞與兩個詞之間的關係

分散式表示🡪向量空間的詞彙表示(上下文預測當前詞義或當前詞預測上下文的詞)🡪將詞彙轉換成低緯度稠密向量(大致上可分為兩種:語法關係、語義關係)🡪類比運算(KING:KINGS = QUEEN : QUEENS / 美國:華盛頓 = 台灣 : 臺北)🡪向量運算(令a =MAN , a\* = WOMAN , b = KING b\*= QUEEN) 再經由多種可能的cos運算得出之間的關係

**問答系統應用**

利用史丹佛的相依剖析器得出一個結果🡪句子相似性計算(嵌入層🡪編碼層🡪比較層 以向量的方式進入，在比較層依照關聯性給予一個”相關分數”)

**機器翻譯應用**

Google多語言神經機器翻譯系統將**日文**翻譯**韓文**(日文**↔**英文 韓文**↔**英文 日文🡪英文🡪韓文)

**意見探勘應用**

希望從冷冷的文句當中分析出情緒(正面/負面/中性)🡪台大情緒辭典🡪各大領域有不同的名詞被用來當形容詞比喻好壞、正反面🡪仍是一個挑戰

**結論:**

領域知識對自然語言很重要、目前而言資源有限、語言可能包含多種資訊(誇大、諷刺、隱喻…)、還需整合世界知識、人類常識…等多種因素