

中國科學技術信息研究所

植基於詩詞文學特性的詩詞語言
處理及研究之應用

元智大學・中國語文學系

羅鳳珠

gefjulo@saturn.yzu.edu.tw

2009年9月1日

大綱

第一部份：研究動機

第二部份：詩詞文學的特性

第三部份：語意概念標記與分類

第四部分：語意概念分類的應用

第五部分：遭遇的困難與未來的發展

第一部份

- 一、前言
- 二、回顧
- 三、如何讓電腦更接近人腦
- 四、概念體系

前 言 1/2

電腦：具備了比人腦更強的「廣讀群書，博聞強記、搜尋檢索、分析統計」的能力。

但是，電腦只能分辨字形的**形符**，無法解讀字詞含義的**意符**。不能理解語義，不具有思考能力和人的認知、感知、思辨能力。

電腦不具備人的**知識體系**，所以，無法「旁徵博引」與「觸類旁通」；不具備人的**感發能力**，所以無法感受及表達情感；不具備人的**推理思辨能力**，所以無法具有人的哲理思維。

前言 2/2

- 電腦處理有規則資料的能力，以及記憶、搜尋、分析、統計、歸納的能力都遠遠優於人腦。
- 在可預見的未來，在文學研究的領域，電腦難以取代人腦，難以代替人去做所有的事。
- 如何讓電腦做得比人好的部分，交給電腦去做，人所節省下來的時間，去做電腦還無法做到的事，並設法教電腦學會人所具備的能力。
- 如何讓電腦具備人的知識體系以及情意感發與哲理思辨能力，是文學數位化研究未來發展的重點。

1987年，筆者應《國文天地》編輯的邀請，以《探一探文史資料自動化的路》為題，訪問資訊科學領域張仲陶教授，文史領域周何教授、毛漢光教授、王邦雄教授、王熙元教授等五位學者，從資訊科學及經、史、子、集不同領域的角度，提出他們對文史古籍數位化的看法，並從中探索發展的方向。

- 張仲陶教授提出：「不要問電腦能做什么？而是問你要電腦做什么？」的看法，主張由文史界負責提出需要，由資訊界負責滿足需要：「電腦做得比人好的，交給電腦做，人去作電腦還做不到的事。」（資訊科技）
- 周何教授認為：「電腦是很呆板的東西，但怎樣使他具有高層次的功能，幫助人腦體會，這是我所期望的。」（經）

回顧

3/6

- 毛漢光教授參與第一期史籍自動化計畫（1984年7月），負責《食貨志》資料分析、規劃、系統分析，比其他幾位教授，多了實際參與文史資料數位化的實務經驗。毛教授說：「就個人經驗言，在文史自動化的過程中，成敗的關鍵在文史界，不在電腦界。依我的經驗，個人想到的，只要分析出來，電腦都可以做到。所以文史自動化成敗的關鍵在文史界，不在電腦界。」（史）
- 王邦雄教授也認為：「文史自動化不能失去人的主導地位。電腦畢竟不是人，無法做創發性的工作。」（子）
- 王熙元教授認為：「研究工作最重要的是資料的運用，將研究資料經過分析、整理、歸納，分門別類建立資料庫，才能符合文史研究的需要。」（集）

- 臺灣中央研究院於1984年7月開始的「史籍自動化計畫」，除了參與該計畫的張仲陶、毛漢光二位教授之外，其餘三位教授未曾使用過數位資料庫，在當時也不會用電腦。
- 彙整五位學者所提出的觀點，有三項共同處，其一為：電腦不能取代人腦；其二為：引用電腦節省處理資料的時間，人可以做更多思考性、創發性的工作；其三為：古籍數位化成敗的關鍵在文史學界，不在電腦學界，需要由文史學界提出需要，電腦學界滿足需要，二者的通力合作才是古籍數位化成功的關鍵。
- 而今時隔二十多年，張仲陶、周何、王熙元三位教授均已作古多年，他們所提出來的觀點，仍然是文史學界與資訊學界努力的方向。

- 1996年在元智大學講授詩詞課程，每每苦於詩詞習作批改的負擔，突然體會到張仲陶教授所說「電腦做得比人好的，交給電腦做，人去作電腦還做不到的事。」
- 今人作詩填詞最大的困難是無法熟記詩詞格律譜以及平仄聲調、韻目韻字，而這些都是固定的，正是電腦可以做得比人好的部分，於是申請經費，開發詩詞格律自動檢測索引教學系統。

回 顧

6/6

- 1997年8月赴北京大學訪問，得有機會持此系統向袁行霈教授請教，袁教授問我，**這個系統能將用錯的字改成對的，能不能將用得不好的字改得更好？**（倚聲填詞二、二）
- 誠如袁教授所說的，這個系統能將用錯的字改成對的字（完全符合平仄押韻），但是不但不能改成語義情境更好的字詞，甚至於在作自動檢測時，對於一首平仄格律完全正確，但是文句語義全然不通的作品束手無策，**關鍵在於電腦不具有理解語義的能力，更不具有思考、感發的能力，甚至於連面對破音字的平仄判斷能力都沒有，而人腦具備這些能力。**

如何讓電腦更接近人腦

1/4

- 如果讓電腦具備理解語義的能力，就必須讓電腦的運算更接近人腦的思維過程，需要建立符合人類知識體系的語義概念資料庫，使電腦初步具備人的知識概念體系（Ontology）。
- 1945年麻省理工學院副校長Vannevar Bush構思著建立一個儲存及檢索的機器，能儲備所有文獻，檢索時能符合人類的思維邏輯，建立知識連結的軌跡以進行聯想索引，需要引用時，所有的資料都可以隨意檢索出來，並且可以快速的找到另一個相關的資料。
- Bush嘗試要解決的便是以電腦幫助及代替人腦做某些事。

如何讓電腦更接近人腦

2/4

- 1956年，「**電腦與人腦比較**」的議題開始受到關注，美籍匈牙利科學家John von Neumann 應耶魯大學邀請演講，演講的主題是電腦與人腦，並與合作夥伴開始研究建構一個與活腦極為相似的人造機器的可能性，**但是由於對「人腦的語言」**（language of the human brain）還是很不清楚，以致于並未成功。

- 1970年代美國哲學家弗德（J. Fodor）提出了「思想語言的假設」（Hypothesis of Language of Thought），引起廣泛的討論。
- 翻譯《電腦與人腦》一書的譯者蔡耀明在譯者序文《電腦科學的瞻前顧後》指出：「人腦（心智）的語言到底是如何，仍是未定之數。在方興未艾的認知科學，以結合哲學、語言學、認知心理學、腦神經科學、人工智慧等跨學科共同研究的努力下，相信總有一天會解開人類心智的奧秘，屆時要實現以人類心智為模型的電腦，就不再是難事了。」

（引自約翰·馮諾曼（John von Neumann）著，蔡耀明譯，《電腦與人腦》，臺灣商務印書館，2000年初版，〈電腦科學的瞻前顧後〉（譯者序），第7頁。）

如何讓電腦更接近人腦

4/4

- 以語意標記使電腦從分辨字形的能力提升到分辨字義的能力。
- 以語意概念分類幫助電腦理解人的知識體系，具備認知的能力。
- 以情感計算（Affective Computing）及情境感知（Context Awareness）提升電腦情感識別與情感表達的能力，使電腦具備感知的能力。

概念體系 1/6

George Lakoff:

- 何謂概念體系？他是如何組成的？
 - 所有的人是否都有同樣的概念系統？
 - 人的客觀認知與主觀感知的區別
 - 人的肉體經驗和運用想像機制的方式，對於人如何建構範疇和理解經驗是重要的。
- ~Wormen, Fire, and Dangerous Things: What Categories Reveal about The Mind~

概念體系 3/6

人的客觀認知與主觀感知

人的「客觀認知」+「主觀感知」+「思維體系」=構成「概念系統」

人以生理經驗（感官系統：耳鼻口眼皮膚心智）感知客觀外物。

因為經驗、情感、性格的不同而產生不同的感知。

客觀事物透過想像的機制而創造不同的主觀感知。

客觀認知是理性的，個別差異比較少

主觀感知是感性的，個別差異比較多

概念體系^{4/6}

所有的人是否都有同樣的概念系統？

人以生理經驗（感官系統：耳鼻口眼皮膚心智）感知客觀外物，所產生的概念體系，有共通性（多數人相同）。

但是因某些個別的差異而有例外。

- 例如：「花月人酒」被多數詩人認為是自然界與人世間美好的事物：
- 李白：「花間一壺酒，獨酌無相親。舉杯邀明月，對影成三人。」
- 李商隱：「縱使有花兼有月，可堪無酒又無人。」
- 李煜：「春花秋月何時了」（亡國的經驗）

概念體系 5/6

人的概念系統

如果人的概念系統是有規則的

如果所有的人都有同樣的概念系統

如果多數人的多數概念是相同的

那麼，以本體知識解構文學作品，重新建構
成為電腦可以讀懂的知識體系，便可以使
電腦具備人的知識體系。

概念體系 6/6

本體論

用本體知識的基本元素（概念）描述知識表達的結構（概念之間彼此的關聯）。

知識本體論是用於描述及表達特定領域知識的一組概念或術語，它可以用來研究特定領域知識的類別、類別的屬性，以及類別與類別間的關係，進一步達成概念語意的資訊檢索。

第二部份：詩詞文學的特性

一、詩詞文學的內容

二、文學的表達與共鳴

三、詩詞文學語言的特殊性

詩詞文學的內容

文學的內容

- 一、人與外物的互動
- 二、客觀的外在景物
- 三、主觀內在的心靈活動

以賀鑄，〈眼兒媚〉為例：

蕭蕭江上荻花秋，作弄許多愁。（人與外物的互動）

半竿落日，兩行新雁，一葉扁舟。（客觀的外在景物）

惜分長怕君先去，直待醉時休。今宵眼底，明朝心上，後日
眉頭。（主觀內在心靈的活動）

詩詞文學的內容：元素^{1/3}

詩詞文學內容的元素：人、事、時、地、物
那些人在那些時間，那些地方做過那些事，
牽涉那些物。

小說、戲曲等多數文學作品的內容都包含這
五項元素，而且都是直接描寫。

散文、詩詞因篇幅較小，其表述方式，或直
接描寫，或隱藏在字裡行間，間接表述。

詩詞文學的內容：元素^{2/3}

直接描寫：

以「柳宗元〈江雪〉：千山鳥飛絕，萬徑人蹤滅。孤舟蓑笠翁，獨釣寒江雪。」為例

人：蓑笠翁

事：獨釣

時：寒冬

地：江邊

物：孤舟、鳥

詩詞文學的內容：元素^{3/3}

間接表述：

雖未直接描寫，仍於詩意之中表達

以杜甫〈八陣圖〉為例：

「功蓋三分國，名成八陣圖。江流石不轉，遺恨失吞吳。」

人：諸葛亮、劉備、孫權、周瑜、曹操

事：吳、蜀、魏三分天下之爭戰，劉備遺恨失吞吳

時：三國時代

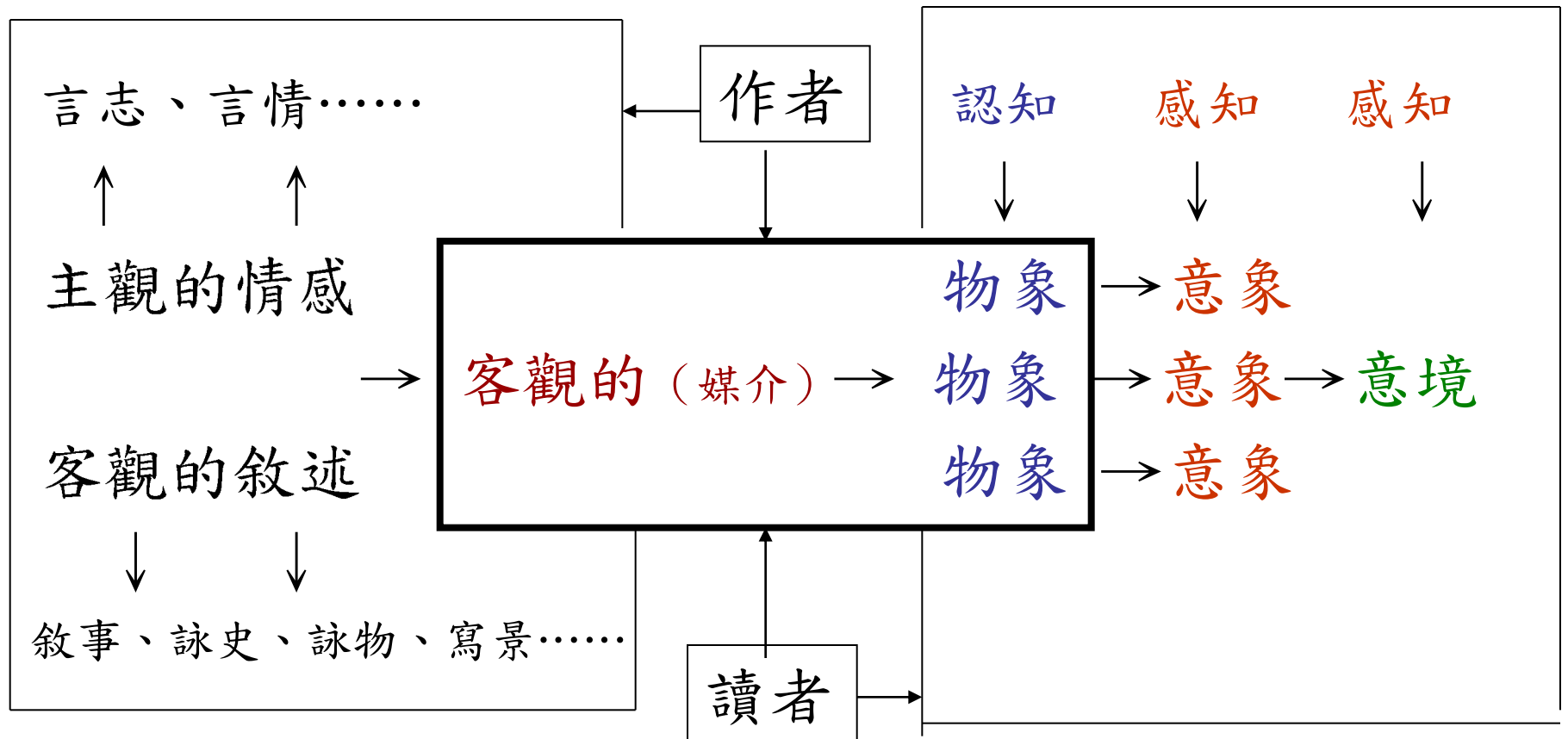
地：魚復（腹）＝今四川奉節縣

物：八陣圖

文學的表達與共鳴：以客觀物象引起共鳴

一、形式：書寫文字、使用文體

二、內容：言志、言情、敘事、詠史、詠物、寫景……



文學的表達與共鳴：以客觀物象引起共鳴

賀鑄〈青玉案〉：試問閒愁都幾許？一川湮草，滿城風絮，梅子黃時雨

作者抽象的情感：試問閒愁都幾許？

閒愁都幾許？如何量長短、稱輕重、衡多寡、掂厚薄？

作者與讀者共有的經驗：一川湮草，滿城風絮，梅子黃時雨
（都是具體可見可知可感的物象） → 引發意象

以具體的物象作為引發作者與讀者共鳴的媒介

媒介：文字 → 物象 → 意象

文學的表達與共鳴：以**感官經驗**引起共鳴

人以**感官感受**外物：作者透過感官感受的描寫以引發讀者相同的**感官經驗**，進而體會作品，與作者共鳴。

李煜〈玉樓春〉

晚妝初了明肌雪。春殿嬪娥魚貫列（**視覺**）。鳳簫吹斷水雲間，重按霓裳歌遍徹。（**聽覺**）臨風誰更飄香屑（**嗅覺**）。醉拍闌干情味切（**味覺**）。歸時休放燭花紅，待踏馬蹄清夜月（**心境情境綜合感覺：情意、情境感知**）。

蘇軾〈永遇樂〉：明月如霜，好風如水（**觸覺**）。

詩詞文學語言的特殊性1/6：構詞

構詞的特殊性

除了使用專有名詞或典故之外，以雙音節及單音節詞為主，尤以雙音節詞居多。

因遷就雙音節及平仄格律，使得詩詞語言的構詞方式特殊。

有些詞彙只用於詩詞，不用於其他文體。或該詞彙在詩詞裡是通順而有意義的詞彙，但是當抽離該句詩詞時，這個詞彙往往不通順或無法顯示完整語義。

詩詞文學語言的特殊性2/6：構詞

1. 散文少用：

白居易：「**離離**原上草，一歲一枯榮」、蘇武：「參辰皆已沒，**去去**從此辭」、柳永：「念**去去**，千里煙波，暮靄沉沉楚天闊。」

2. 分散在上下句：

蘇軾：「詩人老去**鶯鶯**在，公子歸來**燕燕**忙」

3. 縮略倒裝語

蘇軾：「如今勝事無人共，花下**壺盧**鳥勸**提**（花下**提壺****盧**鳥勸人提起**葫蘆**喝酒）」

4. 名詞雙舉並列，以豐富詞義：

孟浩然：「把酒話**桑麻**」、蘇庠：「寄語故時**猿鶴**侶」

詩詞文學語言的特殊性3/6：語意

一般用語：只使用單一語義

所使用的詞彙含義與其他文體無異。

無論該詞彙含有幾種含義，「一般用語」一次只使用其中一種含義。

例如：

杜甫〈曲江〉：穿花蛺蝶深深見，點水蜻蜓款款飛。

李白〈登金陵鳳凰臺〉：三山半落青天外，二水中分白鷺洲。

詩詞文學語言的特殊性4/6：語意

特殊用語：雙關義、引申義、典故、象徵、比喻……，同時使用多元語義

詞義的特殊：使用雙關義、引申義以豐富詩意。（其他文體講究語意精確，詩詞正好相反）

韓翃〈章台柳〉：章台柳，章台柳，昔日青青今在否？縱使長條似舊垂，也應攀折他人手。

柳姬〈楊柳枝〉：楊柳枝，芳菲節，苦恨年年贈離別。一葉隨風忽報秋，縱使君來豈堪折。

「柳＝柳樹＝柳姬」（雙關義）＝離別折柳相贈＝柳姬身陷風塵，只如柳葉，終日送往迎來（引申義）

以文字多義或賦予雙關義、引申義的特性來豐富內容。

詩詞文學語言的特殊性5/6：語意

特殊用語：雙關義、引申義、典故、象徵、比喻……，同時使用多元語義

詞義的特殊：使用典故義

蘇軾：「共誰交臂論今古，只有閑心對此君。」（此君＝酒）

蘇軾：「若對此君仍大嚼，世間那有揚州鶴。」（此君＝竹子）

詞義的特殊：深層義

杜甫〈江南逢李龜年〉：「岐王宅裡尋常見，崔九堂前幾度聞。
正是江南好風景，落花時節又逢君。」

落花時節＝植物的落花時節＋人生的晚年＋國運的衰微期，有三層含義。

詩詞文學語言的特殊性6/6：語意

情境用語：詞義特殊，語義須與其他詞彙共現才能顯示

以「苔蘚」為例

- 苔蘚植物是由水生生活趨向陸生生活的植物族群，生長於極端潮濕的向陽地區或是陰暗的環境中，以配合其親水性結構，不然便很快枯死。這些地方往往也是人跡罕至之處，這是我們對於「苔蘚」的認知基礎。詩人循此「認知」所引發的「感知」如下：
- 詩詞中寫到苔蘚時，往往是「江邊道路多苔蘚，塵土無由得上衣。」的失意文人，或是「玉階陰苔蒼蒼的苔蘚，君王履綦難再得。」的失寵佳人，或是「雪花天色，羅綺深成的苔蘚斑。」的空閨女子，或是「花木閒門苔蘚生，灑川特去得吟情。」的山林雅士。
- 「苔蘚」若是長於門前或台階，詩人將之與「門」、「階」並列呈現時，更核心的集中描寫因失寵而幽居空閨女子的淒清冷落。鄭谷〈長門怨〉，二首之二：「流水君恩不再使，無人前來，階閒無用而長滿綠苔的蒼涼景象。」

第三部份：語意概念標記與分類

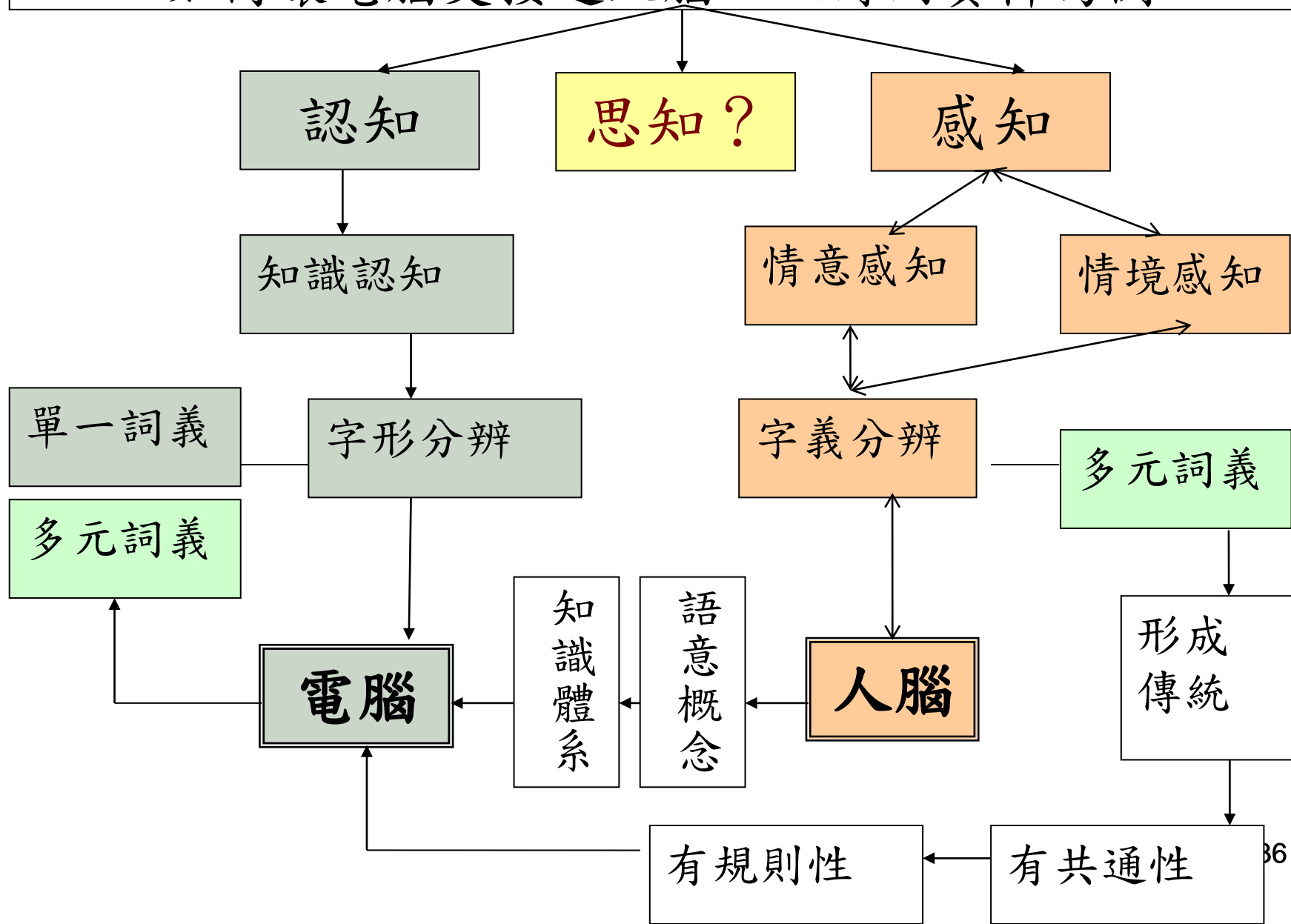
一、詩詞語言的語意概念分類方法

二、詩詞語言的語意概念分類體系

三、本研究語意標記與分類

詩詞語言的語意概念分類方法及分類體系

如何讓電腦更接近人腦：以詩詞賞析為例



詩詞語言的語意概念分類方法及分類體系

- 「人」除了知識認知、情意感知、情境感知之外，還具備由各種單一詞彙的知識所建構的知識體系以及理解感受詩詞作者在詩詞作品內容所創造的豐富語意、語境與情意、情境，這是只能分辨字形的電腦所無法具備的。
- 詞彙經過語意標記，雖然已經能夠使電腦初步具有分辨語意的能力，但也只是零星片段沒有經過組織的語意，不符合人的知識體系，難以觸引對詩詞的感發能力。

詩詞語言的語意概念分類方法及分類體系

本研究以詞彙為單位，先建立詞彙語意標記，再依語意類別建構語意概念，依語意概念分類。

詞彙語意標記→建構語意概念→語意概念分類

電腦除了無法分辨字詞義，也無法判斷詞組的構詞單位，因此，正確的詞彙切分是語意標記重要的關鍵。

詩詞語言的語意概念分類方法及分類體系

詞彙切分^{1/6}

- 本研究以詞彙為單位，先建立詞彙語意標記，再依語意類別建構語意概念，依語意概念分類。
- 詞彙語意標記→語意概念→語意概念分類
- 電腦除了無法分辨字詞義，也無法判斷詞組的構詞單位，因此，正確的詞彙切分是語意標記重要的關鍵。

詩詞語言的語意概念分類方法及分類體系

詞彙切分^{2/6}

- 詩詞的句法有固定的規則，五言詩之前二字，七言詩之前四字，以雙音節詞為主，除非使用專有名詞或典故。
- 唐宋詞除了使用單字領字的句子第一個詞彙是單字詞，其他句子及扣除單字領字之後的句子，也以雙音節詞為主。（對瀟瀟暮雨灑江天）
- 五言詩的句法多數是上二下三，下三部分作上二下一，或作上一下二。
- 七言詩的句法多數是二二三，下三部分作上二下一，或作上一下二。

詩詞語言的語意概念分類方法及分類體系

詞彙切分^{3/6}

依據詩的句法做適切的斷詞，切分的優先順位是：

- 一、專有名詞
- 二、詞的領字
- 三、典故

末三字在名詞及典故優先的前提下，比對已切分出來的詞彙，相符合者，則先認定為詞彙，最後再以人力校正。

詩詞語言的語意概念分類方法及分類體系

詞彙切分^{4/6}

一、建立人名、地名……等各種專有名詞詞彙資料庫（辛棄疾〈玉樓春〉：提壺沽酒已多時，**婆餅焦**時須早去。婆餅焦是飛禽，提胡盧鳥也，杜鵑也。）

二、建立詩詞典故詞彙資料庫

三、建立詩譜、詞譜資料庫，分析其句法，**並將領字位置標示出來。**

詩詞語言的語意概念分類方法及分類體系

詞彙切分^{5/6}

四、詞彙切分方法及步驟：

1. 第一優先比對專有名詞資料庫，以專有名詞做為詞彙切分之最優先順序。
2. 第二優先之切分順序是**領字**：依據所建立的宋詞詞譜優先將領字切分，其餘詞句依「宋詞詞譜」之步驟切分，例如：「**對**／瀟瀟／暮雨／灑／江天／，一番／洗／清秋／。 **漸**／霜風／淒緊／，關河／冷落／，殘照／當樓／。」（只有詞需要此步驟）
3. 第三優先比對典故詞彙資料庫
4. 第四優先比對疊字詞：茫茫、悠悠、灼灼

詩詞之句首，除了使用專有名詞、典故或領字之外，絕大多數都是雙字詞。

詩詞語言的語意概念分類方法及分類體系

詞彙切分 6/6

4. 其餘句子之切分步驟

- 句子字數為偶數句之句子，兩兩切分為雙字詞，例如：「離恨／恰如／春草／，更行／更遠／還生／」。
- 句子字數為奇數句之句子，末三字之外的句子，兩兩切分為雙字詞，例如：「菡萏／香消／翠葉殘／，西風／愁起／綠波間／，還與／韶光／共憔悴／，不堪看／。」。
- 奇數句句末三字，其句法或作上二下一，或作上一下二，並無準則，切分的方法需仰賴詞庫以提高切分之正確率。切分時以前述步驟所獲得之雙字詞比對句末之三字詞，若這三個字之前二字或後二字是前述步驟所獲得之雙字詞，則優先取為雙字詞，予以切分。

不過也有末三字之前二字或後二字都是詞彙的句子，此時便需以人工校正，例如：「歸心正似三春草」（蘇軾〈虞美人〉）之「三春」、「春草」都是宋詞常用的詞彙，遇到這種情況，便需以人工校正。

註：參見拙作，〈[詩詞語言詞彙切分與語意分類標記之系統設計與應用](#)〉，第四屆數位典藏技術研討會，中央研究院主辦，2005年9月1-2日。

本研究語意標記與分類

1/5

一、本研究以《樂府詩集》5,156首、《全唐詩》51,171首、《唐宋詞》24,226首、宋名家詩44,314首、《全台詩》23,066首，**總計147,933首**為範圍，建立詩詞語意標記與分類架構。

二、本研究以詩詞之內容，進行：**詞彙切分** → **語意標記** → **語意概念分類** → **建構語意概念知識分類體系**。

三、語意標記以**詞彙**為單位。

- 予以標記語意及分類的詞彙以實詞為原則，助詞、介詞、連詞、嘆詞等虛詞的語意需附著在實詞上，實詞裡的副詞也須和實詞連結，詞義才能完整顯示，因此這些詞彙暫時不分類。
- 但是虛詞裡的擬聲詞所代表的含義，有些是某些物象的聲音，故予以分類。

本研究語意標記與分類

3/5

- 本研究語意標記與分類以滿足文學教學與研究為主。
- 本研究分類的標的是中國詩詞文獻，所建立的類別是依據詩詞文獻所切分之詞彙，依據前述之原則予以分類，所得結果是詩詞內容所包含的語意，不是完整的人類知識體系，因此有些古典詩詞缺乏的內容，其知識類別便不會出現在詩詞語意概念分類體系裡。

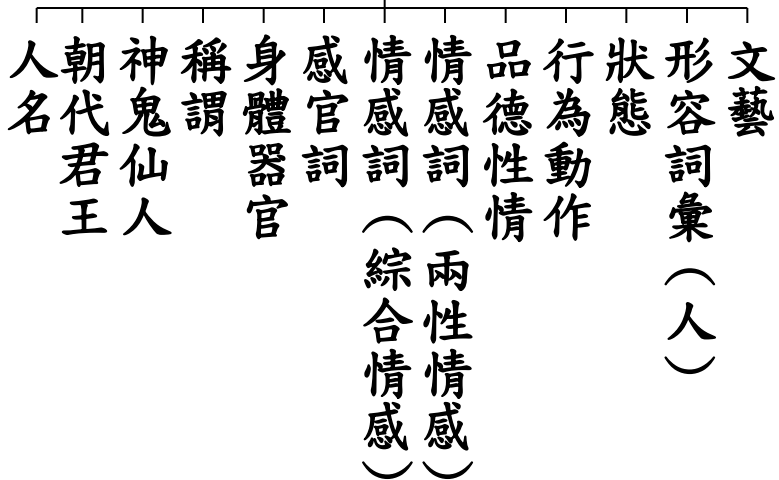
- 本研究將詩詞語意概念分為四層：總類、大類、中類、控制詞彙。
- 本研究因詩詞文體內容所包含的元素，以「人、事、時、地、物」作為最上層的總類，無法包含在這五項元素之外的概念另以「其他」概括之。
- 總計：總類6類，大類54類、中類372類。

所建立的詩詞語意概念分類體系包含二部分：

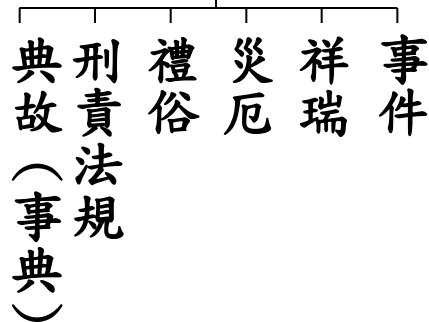
- 其一是以詩詞作品詞彙經過詞彙控制所建立的「**詩詞詞彙表**」。
- 其二是詩詞語意概念分類體系，以「**詞彙索引**」連結「**詩詞詞彙表**」與「**詩詞語意概念分類體系**」二者之間的聯繫關係。
- 關係表如下：

語意概念分類總表

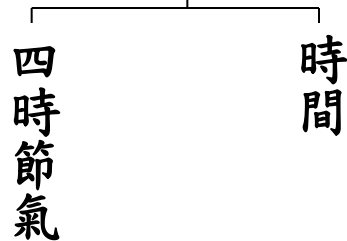
人



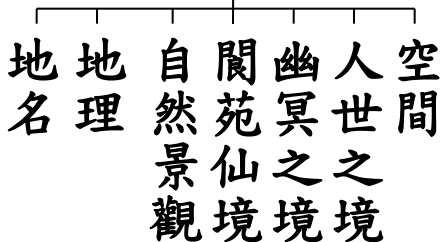
事



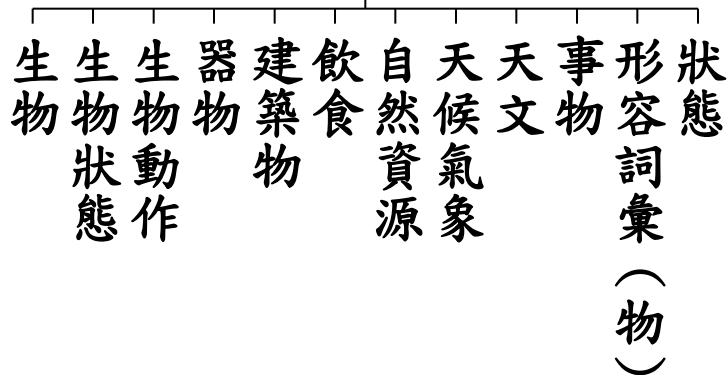
時



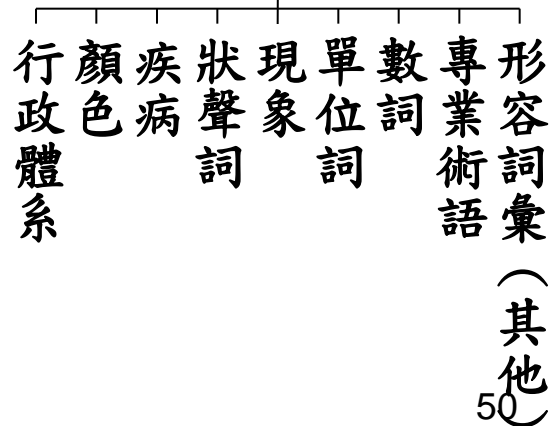
地



物



其他



文藝

書籍泛稱、書籍專名、書篇名、文體、引文、樂曲泛稱、樂曲專名、書畫泛稱、書畫專名、書法、書法（字帖傳名）、舞蹈、戲劇

形容詞彙（人）

氣質內涵、人生機遇、官場境遇

狀態

生理狀態、心智狀態、心理能願、生活狀態、心神氣力、合稱（相反詞、相近詞、相異詞）

行為動作

一般行為（人部、刀部、口部、土部、彳部、日部、水部、糸部、衣部、艸部、言部、貝部、走部、足部、馬部、宀部、支部、走部、其他部）、言語動作、生活活動、官場活動、治國之道、藝文活動、娛樂活動、戰爭活動、行蹤消息

品德性情

正面、負面

情感詞（兩性情感）

傾心愛戀、相思怨懟

情感詞

正面情感（喜悅欣樂、自我肯定、景仰羨慕、悠然自適、憐惜愛寵）、

（綜合情感）

負面情感（自我否定、惱怒怨妒、悲哀傷痛、感懷惆悵、憂愁掛慮、厭煩憎惡、懊惱憾惜、擔驚受怕、辛酸艱苦、慚愧自責、驕矜自恃、思鄉念舊、尷尬為難）、合稱（相反詞、相近詞）

感官詞

嗅覺、聽覺、味覺、視覺、觸覺、綜合感覺詞

身體器官

五官顏面、鬚髮髮膚、身體四肢、五臟器官、分泌物、合稱

稱謂

人稱代名詞、自稱謙稱、尊稱美稱、貶稱惡稱、親人眷屬、指代稱謂、一般稱謂、職業身份、職官爵位、皇室稱謂、宗教稱謂

神鬼仙人

神鬼專稱、神鬼泛稱、菩薩尊號

朝代君王

君王帝號、君王年號、君王陵墓

人名

本名、別名字號、他稱、法號、合稱、外國人名、虛構人名、待考

事

事件

↓ 史事、事情概稱、世局、**國**
家氣運

祥瑞

↓ **福泰祥安、祥瑞預兆**

災厄

↓ 旱災、水災、地震、兵災、
其他災害、災厄預兆

禮俗

↓ 祭祀、婚喪喜慶、其他

刑責法規

↓ 刑責、法規

典故

↓ 佛典、神話典故、其他

時

時間

↓
明確時間、範圍時間（今昔、日、月、世代、生、年、清晨、白晝、黃昏、夜晚、時刻、其他）、相對時間、**年齡時間**

四時節氣 ↓ 節氣、節慶、四季

空間 → 方向、位置、距離

人世之境 → 人間

幽冥之境 → 幽冥

閔苑仙境 → 閔苑、仙境

自然景觀

山峰崖嶺、水澤湖泊、沙洲島嶼、洞穴
坑谷、林原莽漠、山川泛稱、關津渡口
(關隘、津堤、合稱)、風景勝蹟(自然景物、季節景物、人為景物)

地理

全域範圍、境外區域範圍、境內區域範圍
(全域、泛稱、東方、西方、南方、北方)、邦國都城、郊原村野、街道巷弄、特殊場域、場域概稱

地名

行政地名、街坊地名、山嶺地名、河湖地名
(江河溪流、湖海池潭、沙洲)、島嶼地名、
關津地名(關隘、門、津堤、港口、橋)、
虛構地名、外國地名、地名別稱(行政地名、
街坊地名、山嶺地名、河湖地名、島嶼地名、
關津地名)、地名合稱

狀態 → 對比狀態

形容詞彙(物) → 景物形態、植物形態、動物形態、
物品形態、環境形態

事物 → 事物概稱

天文 → 日、月、星、天空、銀河、合稱(同類、異類)

天候氣象 → 風霜雪露、雲霧煙霞、雷電電霓、冷寒、熱
早、晴天、陰雨、合稱(相反詞、相近詞、
相異詞)、概稱

自然資源 → 玉石、火、金屬、礦物、燃料、合稱(同類、
異類)

飲食 → 飲品(茶、酒、其他)、食品、食材、糕點、調
味品、菸草、宴席、合稱

建築物 → 建築專名(宮室屋廬、亭臺樓閣、園林院落、寺廟道
觀、碑塔陵墓)、宮室屋廬、亭臺樓閣、園林院落、
寺廟道觀、碑塔陵墓、公署建築、建築部件、軍事攻

器物 → 生活用品(飲食器具、烹飪器具、裝置容器、起居用
品、其他用品、梳妝用品)、交通工具(水路、陸路、
空路、合稱)、工具用品(兵器單稱、兵器合稱、戰
爭物品、農牧漁獵工、生活工具單稱、生活工具合稱、
刑具)、絲帛服飾(衣冠腰帶、布料、鞋襪、配飾、
合稱)、宗教用品、喪葬祭祀、娛樂器具、金帛錢幣、
公堂器物、旗幟、帳幕、神器、樂器(單稱、合稱、
泛稱)、文具(單稱、合稱、泛稱)、器物合稱、器
物泛稱

生物動作 → 動物動作、植物動作

生物狀態 → 動物生命狀態、植物生命狀態

生物 → 神物(動物、植物、合稱、泛稱)、動物專名(走獸、
飛禽、昆蟲爬蟲、水生兩棲、合稱)、動物指稱(皮
毛、泛稱、其他、合稱)、植物專名(木本、草本、
禾本、苔菌、瓜果、合稱)、植物泛稱(花、木、葉、
枝、果、草、莖、根、苗、合稱)、異類合稱

形容詞彙（其他）

↓
對比詞

專業術語

↓
佛家語、作戰語、法術語、道家語

數詞

↓
定量數詞、概量數詞、等級數詞

單位詞

↓
文藝單位、長度單位、重量單位、面積單位、容量單位、量詞單位

現象

↓
光芒、陰影、自然現象

狀聲詞

↓
自然界聲音、動物聲音、人類聲音、器物聲音

疾病

↓
疾病療法、疾病類別

顏色

↓
原色、借色、間色、顏色概稱

行政體系

↓
部門名稱、朝代名號、朝代名號合稱、
民族邦國名、民族邦國名（合稱）、國名

第四部分：語意概念分類的應用

一. 對資訊檢索質量與使用需求的影響

二. 對詩詞情感表達與識別研究的影響

語意標記及語意概念分類

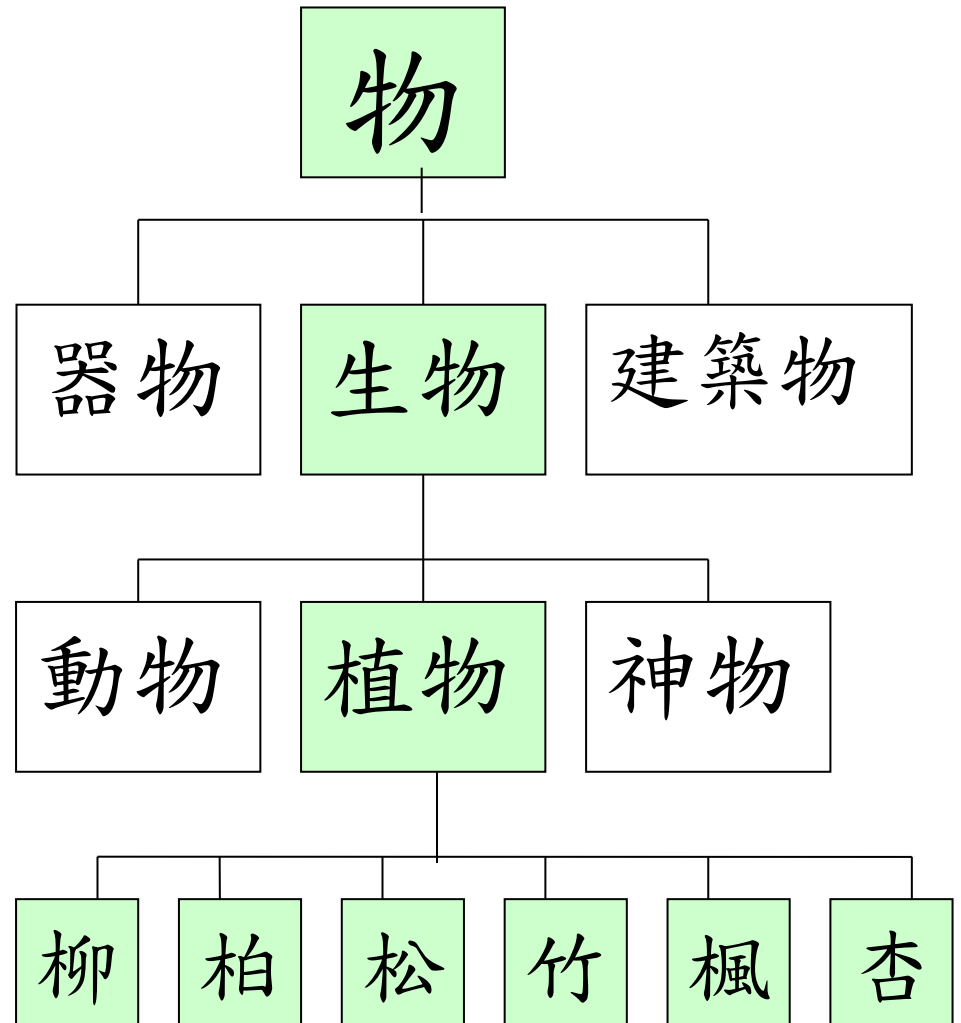
對資訊檢索質量與使用需求的影響

語意概念分為總類、大類、次類、元素四層。

往上層擴大搜尋範圍。

往下層縮小搜尋範圍。

同一層為同一知識領域範圍。



語意概念檢索

有布林功能的欄位可使用「&」=AND、「/」=OR、「-」=NOT

作者：	<input type="text"/>	(含別名字號/布林)
朝代：	請選擇-	
詞牌：	<input type="text"/>	
詞題、序：	<input type="text"/>	(含詞序)
詞文：	<input type="text"/>	(字形/布林檢索)
詞句語意 類別：	總類：請選擇	(測試中)
	大類：請選擇	
	次類：人事	
	元素：時地物其他	
宮調：	請選擇-	
目前只有蘇軾 (1036~1101) 詞含有作品時空標記資料 (1060~1101)		
註解：	<input type="text"/>	
寫作時間：	<input type="text"/>	(紀元年ex.宋仁宗嘉祐五年)
寫作時間：	<input type="text"/>	(西元年ex.1060)
寫作年齡：	<input type="text"/>	
寫及地點：	<input type="text"/>	
寫作地點：	<input type="text"/>	
<input checked="" type="radio"/> 字形檢索		
<input type="button" value="開始檢索"/> <input type="button" value="取消"/>		

語意概念標記應用於全文檢索之驗證：

以唐宋詞網站為例

語意概念標記應用於全文檢索之驗證： 以唐宋詞網站為例（物→天文→星）

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)

← 上一頁 → 搜尋 我的最愛

網址(D) http://cls.hs.yzu.edu.tw/CSP/W_DB/index.htm 移至 連結 >>

語意概念檢索

有布林功能的欄位可使用「&」=AND、「/」=OR、「-」=NOT

作者：	<input type="text"/>	(含別名/字號/布林)
朝代：	請選擇-	
詞牌：	<input type="text"/>	
詞題、序：	<input type="text"/>	(含詞序)
詞文：	<input type="text"/>	(字形/布林檢索)
詞句語意 類別：	總類： 物	(測試中)
	大類： 天文	
	次類： 星	
	元素： 請選擇	
宮調：	請選擇-	
注解：	<input type="text"/>	

目前只有蘇軾

空標記資料 (1060~1101)

三星
三臺
女宿
太白
太白、歲星、辰星、熒惑
太微
太歲
太微

完成 網際網路

以語意概念檢索「星」結果：2076筆

與「月亮」概念相關的同近義詞能被檢索出來，完整性，提高檢索結果的完整性。

共計2076筆

序號	詞牌	作者	朝代	詞題(序)	宮調	詞文	詞中辭彙	辭彙總類	辭彙大類	辭彙次類	辭彙元素
401	阮郎歸	王之道	宋代	二首之一：和廣濟王宰		玉繩低轉斗闌干。	玉繩	物	天文	星	玉繩
402	醉蓬萊	王之道	宋代	代人上高御帶，時在太學		長庚誕貴。	長庚	物	天文	星	長庚
403	好事近	王之道	宋代	何希淵生日		曾見少微初降，	少微	物	天文	星	少微
404	好事近	王之道	宋代	董令升生日		歸/去騎箕尾。	箕尾	物	天文	星	尾星、箕星
405	臨江仙	王之道	宋代	和劉南伯		疏星明耿耿，	疏星	物	天文	星	星

與「星星」概念相關的同近義詞能被檢索出來，提高檢索結果的完整性。

以字形比對檢索「星」：1427筆

http://cls.hs.yzu.edu.tw/CSP/W_DB/index.htm

檔案(E) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H) 連結 >>

【淺顯低唱】若有知音見採—唐宋詞全文資料庫

共計1427筆

... 31 32 33 34 35 36 37 38 39 40 ...

序號	詞牌	作者	朝代	詞題(序)	宮調	詞文
351	浣溪沙	曾覿	宋代	櫻桃		滿林璀璨綴繁星。
352	好事近	王之望	宋代	和榮太監		佳句麗星日。
353	減字木蘭花	王之望	宋代	恭人生日		與爾尊前作壽星。
354	醉花陰	王之望	宋代	生日		休畫老人星，
355	風流子	王之望	宋代	范覺民生日		中有麗天星斗，
356	錦堂春	葛立方	宋代	正旦作		星翁說道宜官。
357	滿庭芳	葛立方	宋代	五姪將赴當塗，自金壇來別		別來幾換星霜。
358	水調歌頭	王識	宋代	觀星		初出極星未遠，
359	水調歌頭	毛并	宋代	上元郡集		十里寒星相照，
360	漁家傲引	洪适	宋代	伏以黃童白叟，皆是煙波之釣徒；青笠綠蓑，不識衣冠之盛事。長浮家而醉月，更轆棹以吟風。樂哉生涯，翻在樂府。相煩女伴，漁父分行。		碧落雲高星爛爛。

... 31 32 33 34 35 36 37 38 39 40 ...

與「星星」概念無關的「壽星」被檢索出來，
正確性不足。

語意概念標記應用於全文檢索之驗證： 以唐宋詞網站為例（物→天文→月）

http://cls.hs.yzu.edu.tw/CSP/W_DB/index.htm

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H) 連結 >>

【淺顯低唱】若有知音見採—唐宋詞全文資料庫

語意概念檢索

有布林功能的欄位可使用「&」=AND、「|」=OR、「-」=NOT

作者：	<input type="text"/>	(含別名字號/布林)
朝代：	請選擇-	
詞牌：	<input type="text"/>	
詞題、序：	<input type="text"/>	(含詞序)
詞文：	<input type="text"/>	(字形/布林檢索)
詞句語意 類別：	總類： 物	(測試中)
	大類： 天文	
	次類： 請選擇	
	元素： 請選擇	
宮調：	請選擇-	
目前只有蘇軾		空標記資料 (1060~1101)
註解：	<input type="text"/>	
寫作時間：	<input type="text"/>	(紀元年ex:宋仁宗嘉祐五年)
寫作時間：	<input type="text"/>	(西元年ex:1060)
寫作年齡：	<input type="text"/>	
寫及地點：	<input type="text"/>	
寫作地點：	<input type="text"/>	

Dropdown menu options: 日, 月, 合稱(同類), 合稱(異類), 星, 銀河

以語意概念檢索「月」結果：6734筆

http://cls.hs.yzu.edu.tw/CSP/W_DB/index.htm

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H) 連結 >>

【淺顯低唱】若有知音見採一唐宋詞全文資料庫

回首頁 文字檢索 網路展書

回檢索畫面

共計6734筆

... 11 12 13 14 15 16 17 18 19 20 ...

序號	詞牌	作者	朝代	詞題(序)	宮調	詞文	詞中辭彙	辭彙總類	辭彙大類	辭彙次類	辭彙元素
121	定西番	張先	宋代			小逐畫船風月，	月	物	天文	月	月
122	定西番	張先	宋代	執胡琴者九人		競／嬋娟。	嬋娟	物	天文	月	嬋娟
123	少年遊慢	張先	宋代			梯漢同登蟾窟。	蟾窟	物	天文	月	蟾窟
124	翦牡丹	張先	宋代	舟中間雙琵琶		玉盤大小亂珠迸。	玉盤	物	天文	月	玉盤

與「月亮」概念相關的同近義詞能被檢索出來，提高檢索結果的完整性。

以字形比對檢索「月」檢索：7705筆

與「月亮」概念無關的時間詞「月」被檢索出來，正確性不足。

http://cls.hs.yzu.edu.tw/CSP/W_DB/index.htm

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)

【淺甄低唱】若有知音見採—唐宋詞全文資料庫

回首頁 文字檢索 網路展書

回檢索畫面

共計7705筆

序號	詞牌	作者	朝代	詞題(序)	宮調	詞文
161	謁金門	晏殊	宋代			風月宛然無異。
162	破陣子	晏殊	宋代			人間好月長圓。
163	浣溪沙	晏殊	宋代			月好謾成孤枕夢，
164	浣溪沙	晏殊	宋代			三月和風滿上林。
165	浣溪沙	晏殊	宋代			鬢禪欲迎眉際月，
166	鵲踏枝	晏殊	宋代			明月不諳離恨苦。
167	紅窗聽	晏殊	宋代			風前月底，
168	採桑子	晏殊	宋代			陽和二月芳菲遍，
169	採桑子	晏殊	宋代			淡月朧明。
170	喜遷鶯	晏殊	宋代			堯莫隨月欲團圓。

語意概念標記應用於全文檢索之驗證： 以唐宋詞網站為例

- 一、以「星」、「月」為例
- 二、比較以「字形比對」及「語意概念」檢索結果

	語意概念檢索	字形比對檢索
星	2076筆	1427筆
月	6734筆	7705筆

第四部分：語意概念分類的應用

一. 對資訊檢索質量與使用需求的影響

二. 對詩詞情感表達與識別研究的影響

研究方法與研究範圍

1/4

作者/項目	文體別	詞作風格	作品數量	詞彙數量
蘇軾	詩		2,855	95,297
蘇軾	詞	豪放	379	12,418
辛棄疾	詞	豪放	629	26,608
溫庭筠	詞	婉約	74	1,685
柳永	詞	婉約	218	10,670
秦觀	詞	婉約	148	5,815
李清照	詞	婉約	53	1,783
總計			4,356	154,276 ₆₈

研究方法與研究範圍 2/4

蘇軾詩、溫庭筠詞、柳永詞、蘇軾詞、秦觀詞、李清照詞、辛棄疾詞

不同時代背景	晚唐：溫庭筠； 北宋：柳永、蘇軾、秦觀； 南宋：李清照、辛棄疾
不同作者的不同文體	詩：蘇軾； 詞：溫庭筠、柳永、蘇軾、秦觀、李清照、辛棄疾
同一作者的不同文體	蘇軾詩、蘇軾詞
同一文體的不同作者	詞：溫庭筠、柳永、蘇軾、秦觀、李清照、辛棄疾
相同文體的不同風格	豪放詞：蘇軾、辛棄疾； 婉約詞：溫庭筠、柳永、秦觀、李清照

研究方法與研究範圍 3/4

蘇軾詩、溫庭筠詞、柳永詞、蘇軾詞、秦觀詞、李清照詞、辛棄疾詞

作者窮達順逆境遇懸殊	蘇軾、李清照、辛棄疾
性別不同	李清照為女性。 性別不同，生活經驗與情感感受、表達方式不同
政治環境背景不同	承平：柳永、蘇軾、秦觀； 戰亂：李清照、辛棄疾
時代盛衰不同	盛：柳永、蘇軾、秦觀； 衰：溫庭筠、李清照、辛棄疾
文體發展階段不同	初期：溫庭筠； 中期：柳永、蘇軾、秦觀、李清照； 晚期：辛棄疾

研究方法與研究範圍 4/4

- 除了以所選研究範圍的七家詩詞分析比較，為了求得論述的客觀性，另外加上三種數據相互比較，其一是**不含蘇軾詩之六家詞**綜合數據；其二是**七家詩詞**之綜合數據；其三是**唐宋詞全部作品**的綜合數據。
- 本研究進一步透過情感計算方法，分析不同文體、不同作者、不同風格、不同境遇的詩詞作品之情感表達方式，期以從中歸納詩詞情感表達的方式，並建立詩詞情感表達的模組。

詩詞文學的情感分類及情感計算

情感詞彙分類與計量：情感詞彙 1/4

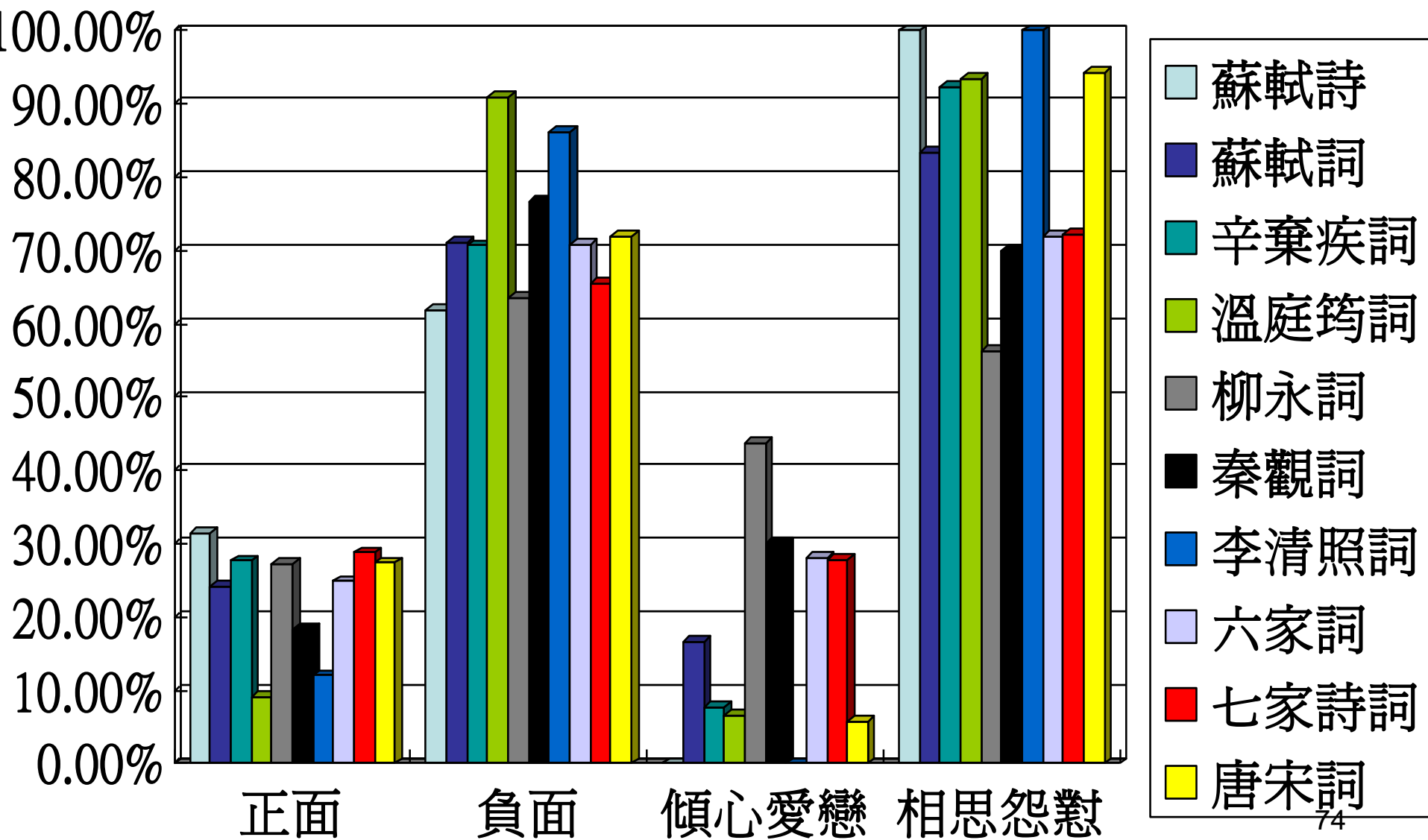
- 本研究依照歷來對人類情感的研究，將情感分為積極正面與消極負面二大類。
- 兩性情感是文學作品的重要內容，但是兩性情感很難以正面或負面區分。
- 本研究將兩性情感獨立，凡是單純書寫男女之情的作品，歸為兩性情感。
- 並將男女之情分為「傾心愛戀」、「相思怨懟」二類，前者偏向正面，後者偏向負面，無法明確劃分或者可以明顯判斷不屬於兩性情感者，歸為綜合情感

詩詞文學的情感分類及情感計算

情感詞彙分類與計量：情感詞彙 2/4

大類	中類	小類之分項	數量
綜合情感	正面情感	憐惜愛寵、景仰羨慕、喜悅欣樂、悠然自適、自我肯定	5類
	負面情感	擔驚受怕、悲哀傷痛、憂愁掛慮、惱怒怨妒、厭煩憎惡、慚愧自責、辛酸艱苦、感懷惆悵、思鄉念舊、驕矜自恃、懊惱憾惜、自我否定	12類
	其他	情感（概稱）、綜合情感（相反合稱）、綜合情感（相近合稱）	3類
兩性情感	傾心愛戀		1類
	相思怨懟		1類

六家七種詩詞的情感詞彙，其情感之不同類別（中類）表列如下 3/4



詩詞文學的情感分類及情感計算

情感詞彙分類與計量：情感詞彙 4/4

- 依據上述分類分析六家七種詩詞的情感詞彙，其情感之不同類別（中類）可知，無論以作者、以文體、以風格區分，詩詞作品描寫負面情感的比例都高於正面情感。
- 再以六家詞之總數及以唐宋詞全文統計，所得到的結果仍然一樣。
- 依此可以推論唐宋詩詞以描寫負面情感的作品居多。

詩詞文學的情感分類及情感計算

相關情境分類與計量：季節 1/10

- 人的情感，會受到哪些自身情緒之外的因素影響？
- 劉勰《文心雕龍·物色》篇所說的：「春秋代序，陰陽慘舒，物色之動，心亦搖焉。」、「情以物遷，辭以情發」。
- 詩人受外物變遷，心中有所感，發而為詩為詞為文，這是由外而內的心境情感變化；詩人將心中所感，寄託在外物之上，這是由內而外的寄託。

詩詞文學的情感分類及情感計算

相關情境分類與計量：季節 2/10

最容易引發詩人感於外物，首當其衝的是季節更迭所引起的天氣、溫度與萬物景觀的變化。

- 傷春與悲秋是中國文學的二十大主題。《毛詩正義·孔穎達疏》曰：「春則女悲，秋則士悲，感其萬物之化，故所以悲也。」
-
- 春秋二季何以最容易引發詩人心中的感發？

詩詞文學的情感分類及情感計算

相關情境分類與計量：季節 3/10

- 懷春、傷春的心情最早見於《詩經》，〈召南·野有死麕〉：「有女懷春，吉士誘之。」
〈豳風·七月〉：「春日遲遲，采繁祁祁，女心傷悲，殆及公子同歸。」寫的都是女子懷春的心情。
- 唐代王昌齡的〈閨怨〉：「閨中少婦不知愁，春日凝妝上翠樓。忽見陌頭楊柳色，悔教夫婿覓封侯。」可以說是為「春女善懷」定了調。

詩詞文學的情感分類及情感計算

相關情境分類與計量：季節 4/10

- 悲秋感慨，源於《楚辭》之〈離騷〉：
「日月忽其不淹兮，春與秋其代序。惟草木之零落兮，恐美人之遲暮。」
- 到了宋玉《九辯》：「悲哉秋之為氣也！蕭瑟兮草木搖落而變衰。」
- 發展至此，深植文人心中悲秋情懷可說已經成熟。

詩詞文學的情感分類及情感計算

相關情境分類與計量：季節 5/10

春女所思所懷，偏向兒女私情，心事難說。
秋士所悲所感，偏向懷才不遇，志業落空。

- 葉嘉瑩認為：「『春女善懷』與『秋士易感』在中國古典詩歌中，其感情是有著相通之處的。寫女子的不被人賞愛和才士的不被人任用，其寂寞悲涼之感，是相似的。」
- 中國詩詞又常以男女關係暗喻君臣關係，以女子不被賞愛暗喻臣子不被重用，如杜荀鶴〈春宮怨〉：「早被蟬娟誤，欲美妝臨鏡慵。承恩不在貌，教妾若為容。」又如王國維〈虞美人〉：「碧苔深鎖長門路，總為蛾眉誤。自來積毀骨能銷，何況真紅，一點臂砂嬌。妾身但使分明在，肯把朱顏悔？從今不復夢承恩，且自簪花，坐賞鏡中人。」
- 無論如何，傷春悲秋已成為中國文學的主要傳統。

詩詞文學的情感分類及情感計算

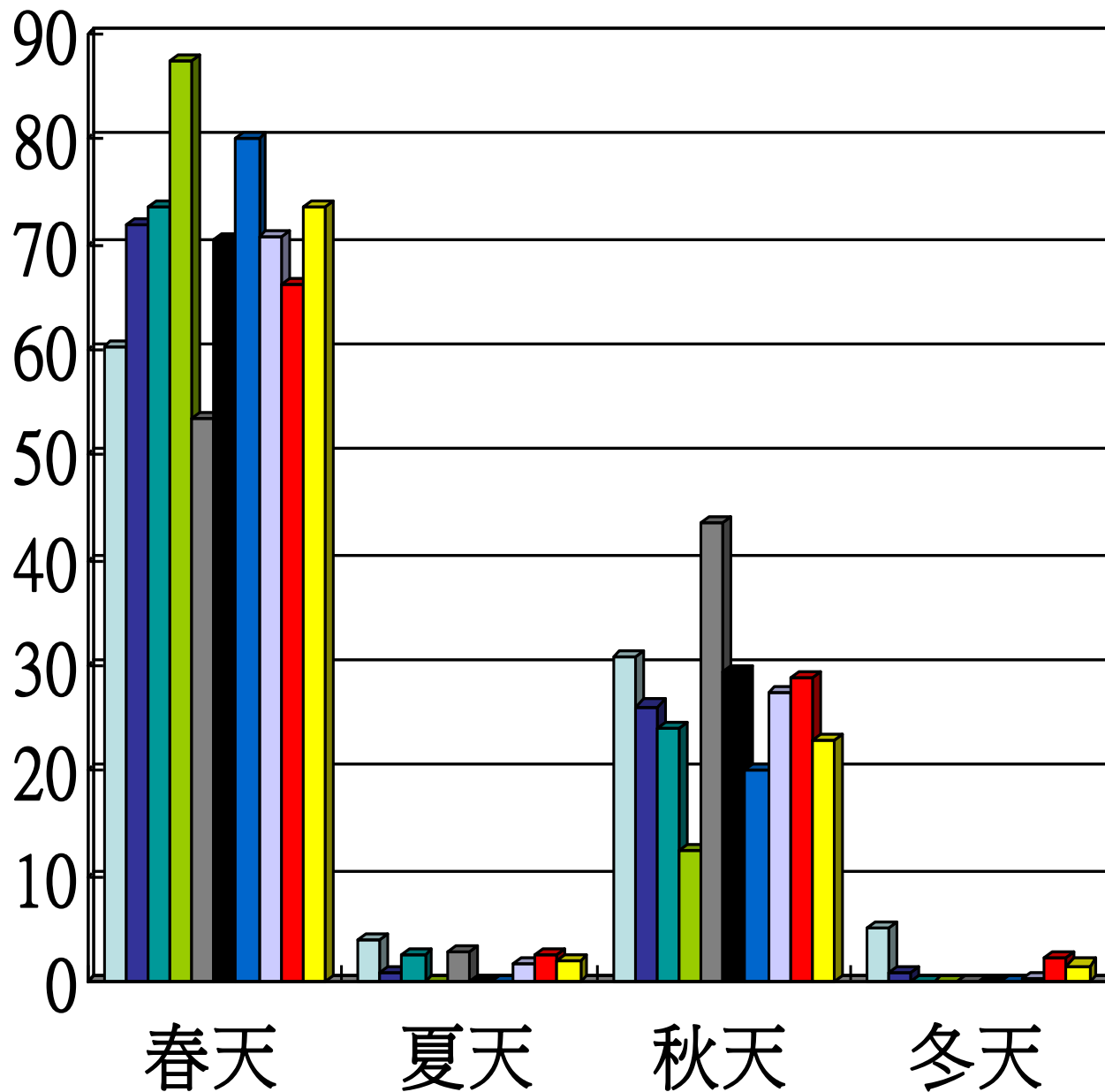
相關情境分類與計量：季節 6/10

柳永詞的內容分為二大部分，一寫男女之情，一寫秋士之感。柳永官場失意的辛酸，使他更有感於宋玉悲秋及「貧士失職而志不平。」的心情，而在詞中直寫「當時宋玉悲感」的悲憤蒼涼，正是「秋士易感」的寫照。

柳永218闕詞作中，引宋玉自況的詞作有五闕，除了〈擊梧桐〉之外，都與宋玉悲秋情懷有關。：

1. 〈戚氏〉：「時宋玉悲感，向此臨水與登山。」
2. 〈玉蝴蝶〉：「望處雨收雲斷，憑闌悄悄，目送秋光。晚景蕭疏，堪動宋玉悲涼。」
3. 〈雪梅香〉：「景蕭索，危樓獨立面晴空。動悲秋情緒，當時宋玉應同。」；
4. 〈爪茉莉〉：「每到秋來，轉添甚況味。…更休道、宋玉多悲」；
5. 〈擊梧桐〉：「見說蘭臺宋玉，多才多藝善詞賦。」

各家作品描寫季節的比例如下表



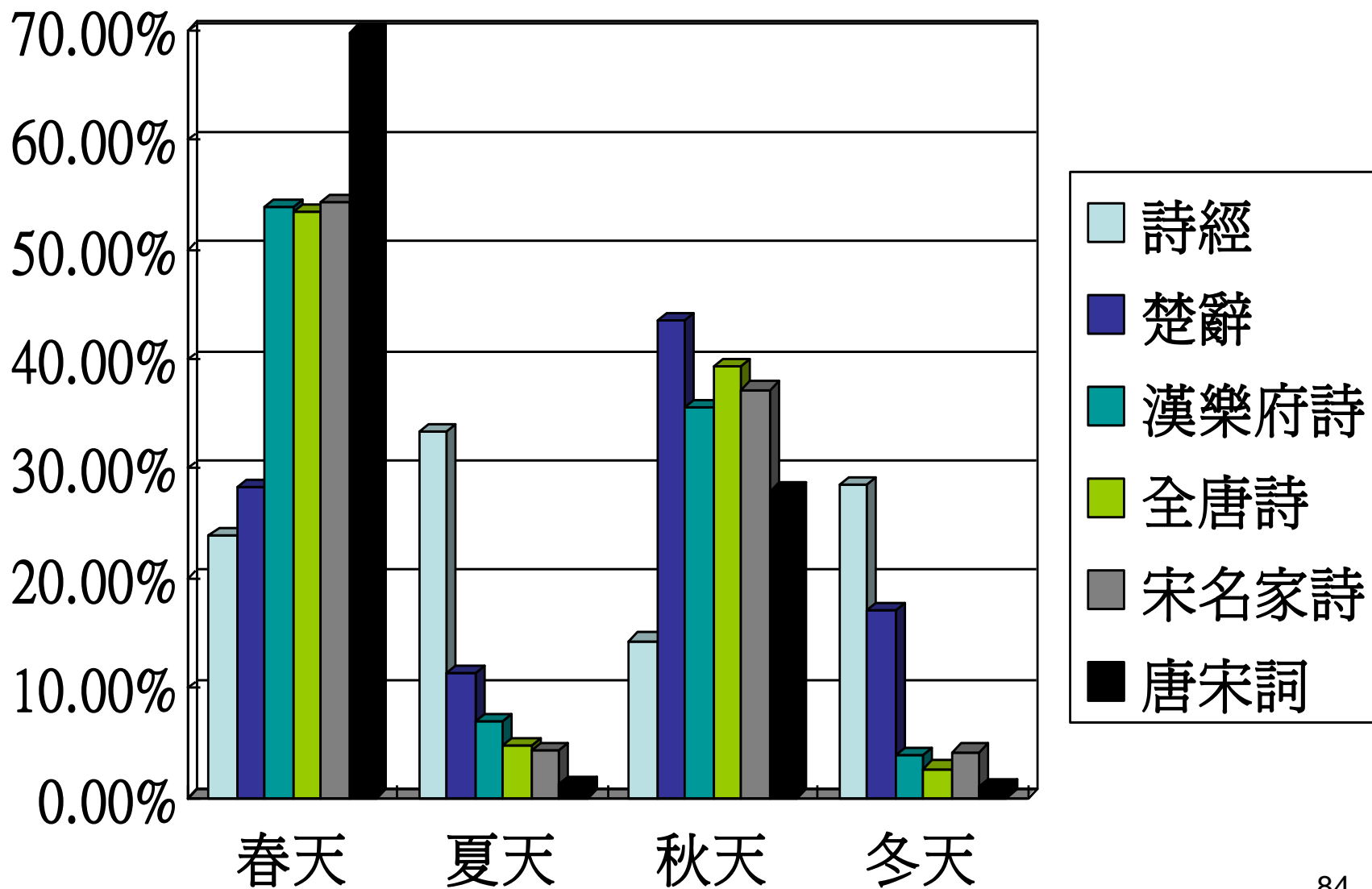
- 蘇軾詩
- 蘇軾詞
- 辛棄疾詞
- 溫庭筠詞
- 柳永詞
- 秦觀詞
- 李清照詞
- 六家詞
- 七家詩詞
- 唐宋詞²

詩詞文學的情感分類及情感計算

相關情境分類與計量：季節 8/10

- 溫庭筠詞的內容，只寫春季的原因，是因為他的詞以「描寫歌妓們的苦痛生活和追求真誠的愛情以及美好光明的願望。」為主，正符合「春女善懷」的論述，秦觀、李清照詞大體類似。
- 柳永描寫秋季的詞雖然比寫春季少，但在七家詩詞中所佔的比例最高，而且高於六家詞、七家詩詞及唐宋詞的綜合比例，是因為柳永仕途失意，深感懷才不遇之悲。
- 秦觀、蘇軾詞寫秋季的比例亦高，因為他們都曾經歷過貶謫（才士不遇）。

宋代以前詩詞的季節描寫比較如下表 9/9



詩詞文學的情感分類及情感計算

相關情境分類與計量：季節 10/10

從上表的分析統計可以看出，《詩經》與《楚辭》的季節描寫比較平均。

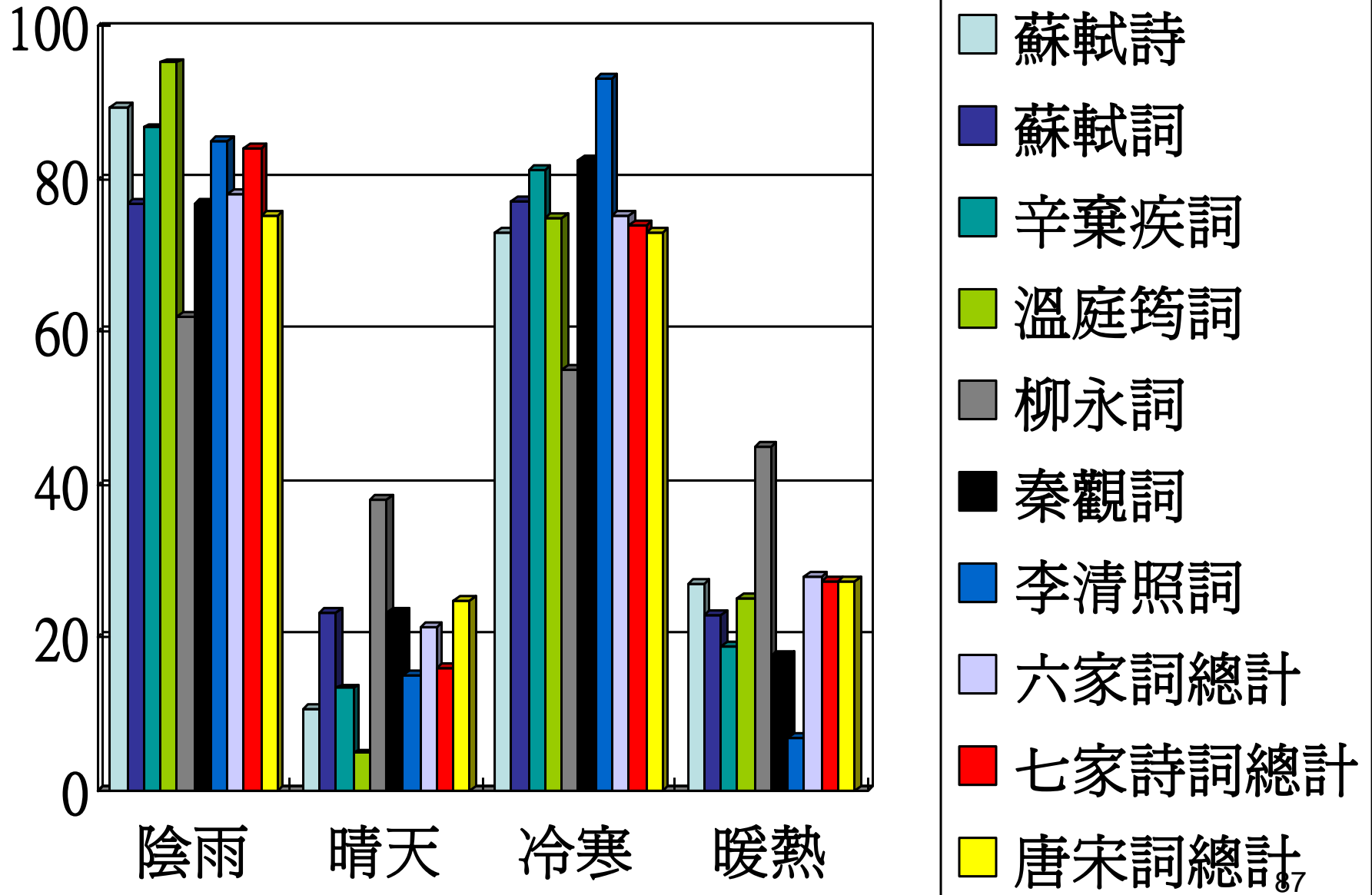
- 但是，在《詩經》的傷春情懷與《楚辭》的悲秋情懷之後的詩詞作品，季節的描寫已經明顯的集中於春秋二季，描寫春天的比例最高，其次是秋天，夏、冬二季的比例非常小，可以驗證出傷春悲秋的情懷已然成為詩詞裡描寫情感的傳統，也是主要的內容。
- 《詩經》與《楚辭》之外，唐宋詩詞作品描寫春天的比例，以唐宋詞最高，描寫秋天的比例以唐宋詞最低，也足以說明以男女之情為主要內容的唐宋詞，正符合春天所思所懷以男女之情為主的傳統。

詩詞文學的情感分類及情感計算

相關情境分類與計量：天候氣溫的變化 1/4

- 隨著季節變化的，除了自然景觀之外，還有天氣陰晴與氣溫暖寒的變化。
- 詩詞既以描寫春秋二季居多，氣溫當然是冷寒多於暖熱。
- 七家詩詞之中，寫陰雨天及冷寒氣溫的比例都高於寫晴天及暖熱天，柳永之外，其他詞人作品二者的比例差距都很大，詳如下表。

各家詩詞描寫天候氣溫的比例 2/4



詩詞文學的情感分類及情感計算

相關情境分類與計量：天候氣溫的變化 3/4

- 冷寒暖熱氣溫與季節相關，這是可以理解的客觀現象。
- 從前面的寫作季節統計，春、秋二季比例高於夏、冬二季，春寒秋涼，所以詩詞寫冷寒多於暖熱，與季節有關。
- 但是天候陰晴與季節的相關性不若溫度這麼高，可是詩詞之中描寫陰雨天的比例遠遠高於晴天，便是一項值得探討的問題。

詩詞文學的情感分類及情感計算

相關情境分類與計量：天候氣溫的變化 4/4

- 詞，屬於南方文學，六家詞人之中，除了蘇軾的詞作有一些寫於比較北邊的密州、開封等地，但寫於南方的比例仍然比較多。
- 中國南方的雨季集中在5-10月，整個中國的平均雨季，夏秋多雨而冬春少雨。
- 詩詞寫於春天的比例高於秋天，所以會遇到雨季的月分是初秋至中秋的9-10月，總計二個月。
- 但是從各個不同範圍的統計顯示，寫陰雨天的比例都遠高於晴天。這是因為詞人筆下的陰雨天，恐怕不全然是天候的陰雨，還包含心境的陰雨。

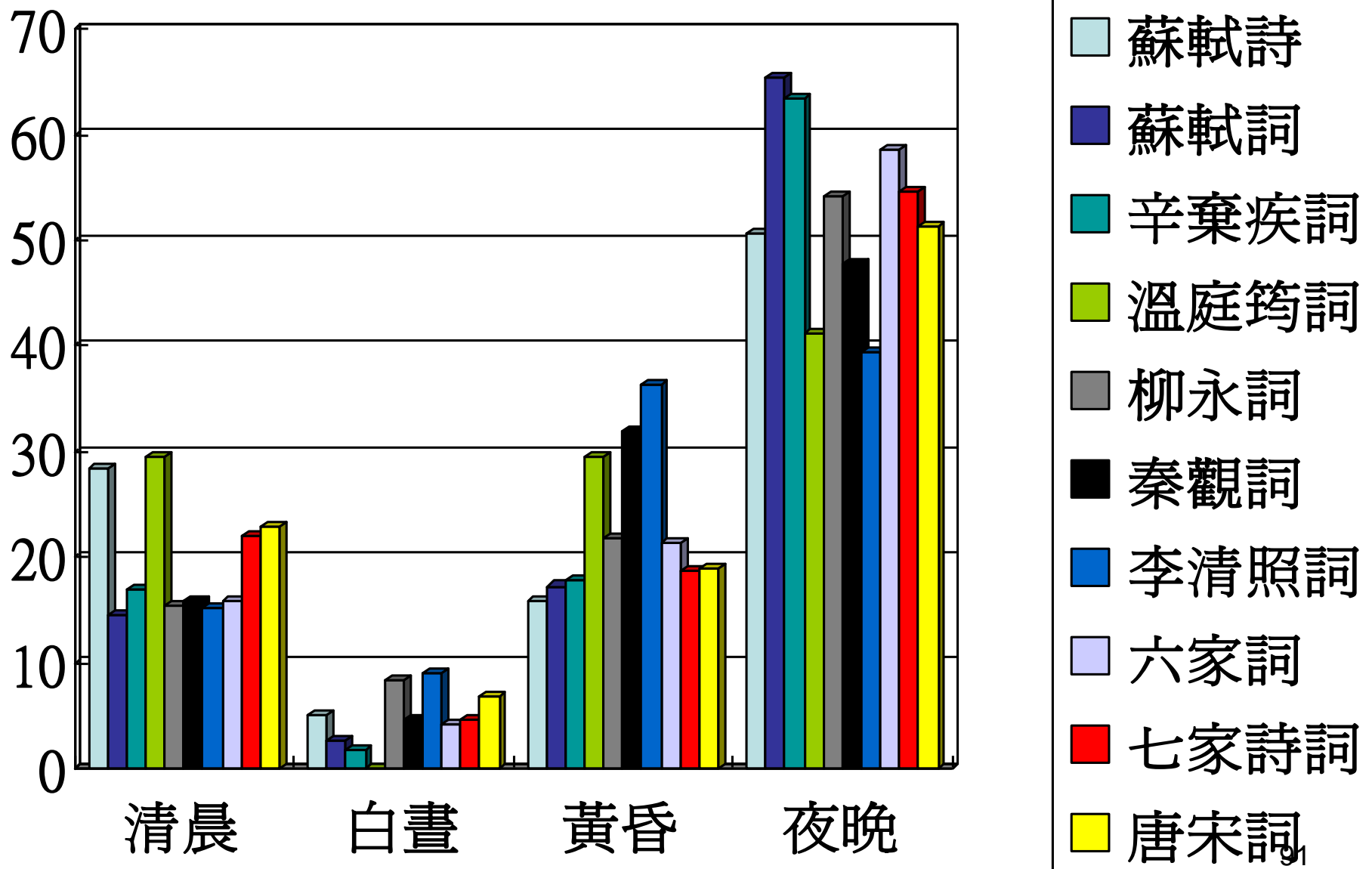
詩詞文學的情感分類及情感計算

相關情境分類與計量：晝夜時間的更迭 1/5

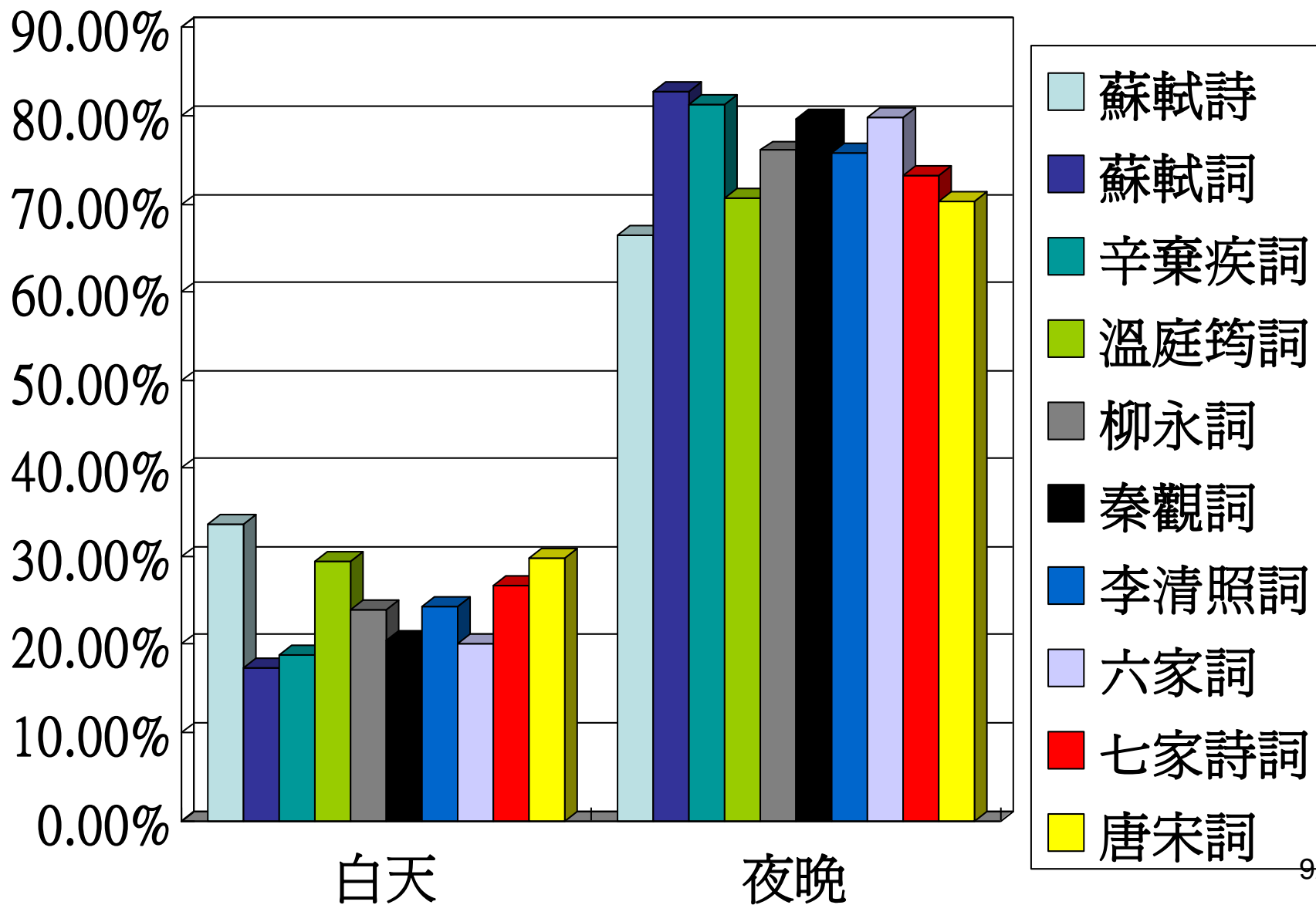
- 另一項影響情感的因素是晝夜所帶來光線明暗的不同，以及黃昏與夜晚特別讓人想家的因素。
- 從下表可知，如果將一天之間的時間分為清晨、白晝、黃昏、夜晚四個時段，所有作品都是寫於夜晚的比例最高，其次多數是寫於黃昏，只有蘇軾詩是寫於清晨的比例高於黃昏，溫庭筠詞寫於清晨與黃昏的比例一樣，比例最低的是寫於白晝。
- 如果將一天之間的時間分為白天（清晨、白晝）與夜晚（黃昏、夜晚）二個時段，則寫於夜晚的比例全部都遠遠高於白天。

晝夜時間光線明暗的變化（清晨、白晝、黃昏、夜晚）

2/5



晝夜時間光線明暗的變化（白天、夜晚） 3/5



詩詞文學的情感分類及情感計算

相關情境分類與計量： 晝夜時間的更迭 4/5

在以負面情感居多的唐宋詞裡，詞人筆下的黃昏：

- 或是「欲黃昏，雨打梨花深閉門」（李煜〈憶王孫〉）的**孤寂**，
- 或是「雨橫風狂三月暮。門掩黃昏，無計留春住。」（歐陽修〈蝶戀花〉）的**無奈**，
- 或是「把酒送春春不語，黃昏卻下瀟瀟雨」（朱淑真〈蝶戀花〉）的**淒清**，
- 或是「已是黃昏獨自愁，更著風和雨」（陸游〈卜算子〉）的**落寞**，
- 或是「小院黃昏人憶別。落紅處處聞啼鴂。」（蘇軾〈蝶戀花〉）的**惆悵**，
- 或是「王孫一去杳無音，斷腸最是黃昏後。」（李之儀〈踏莎行〉）的**傷悲**。

詩詞文學的情感分類及情感計算

相關情境分類與計量：晝夜時間的更迭 5/5

- 黃昏的孤單寂寥，延續到夜晚便成了「夜長衾枕寒」（溫庭筠〈更漏子〉）的**寂寞淒涼**。
- 詞人筆下因「諳盡孤眠滋味」（范仲淹〈御街行〉）而覺長夜漫漫的，**不外乎是空閨女子以及遷客騷人**。
- 顧夐〈訴衷情〉：「永夜拋人何處去，絕來音。香閣掩。眉斂。月將沉。爭忍不相尋。怨孤衾。換我心、為你心。始知相憶深。」**代表了空閨女子的心聲**。
- 李煜〈烏夜啼〉：「昨夜風兼雨，簾幃颯颯秋聲。燭殘漏滴頻敲枕，起坐不能平。世事漫隨流水，算來夢裏浮生。醉鄉路穩宜頻到，此外不堪行。」**寫盡遷客騷人的傷痛**。
- 即便夜裡有一場好夢，夢醒時，恐怕也只落得「夜茫茫，重尋無處，覺來小園行遍」（蘇軾〈永遇樂〉）的**惆悵**。
- **詩詞裡的黃昏夜晚，總是悲傷多，歡樂少，這種情況，也符合了詩詞裡負面情感多於正面情感的現象。**

詩詞文學的情感分類及情感計算

相關情境分類與計量：天文意象 1/8

從詩經〈陳風·月出〉：

月出皎兮，佼人僚（美好）兮，舒窈糾（體態苗條）兮，
勞心悄（憂愁）兮。

月出皓兮，佼人憺（美艷）兮，舒懽受（體態多姿）兮，勞
心忉（不安）兮。

月出照兮，佼人燎（亮麗）兮，舒夭紹（體態輕盈）兮，勞
心慘（悲傷）兮。

月亮已經成了世間多情男女寄託心事的對象。

詩詞文學的情感分類及情感計算

相關情境分類與計量：天文意象 2/8

- 李白〈靜夜思〉：「床前明月光，疑是地上霜。舉頭望明月，低頭思故鄉。」勾起普天之下所有異鄉遊子的思鄉情懷。
- 李白〈關山月〉：「明月出天山，蒼茫雲海間。長風幾萬里。漢下白登道，胡窺青海灣。由來征戰地，不見有人還。戍客望邊色，思歸多苦顏。高樓當此夜，歎息未應閒。」道盡戰地征人與空閨思婦只能遙隔兩地對月嘆息的悲涼。
- 李白〈長相思〉：「長相思，在長安。絡緯秋啼金井欄，微霜淒淒篔簹色寒。孤燈不明思欲絕，卷帷望月空長歎。美人如花隔雲端，上有青冥之長天，下有綠水之波瀾。天長路遠魂飛苦，夢魂不到關山難。長相思，摧心肝。」使多少相隔兩地的有情人一起「長相思，摧心肝」。

詩詞文學的情感分類及情感計算

相關情境分類與計量：天文意象 3/8

- 蘇軾〈水調歌頭〉：「人有悲歡離合，月有陰晴圓缺，此事古難全。但願人長久，千里共嬋娟。」天上的月圓月缺，人間的圓滿缺憾，總是交相更迭，亙古不變。
- 月亮，成了詩人筆下最常被寫及的天文景觀。

詩詞文學的情感分類及情感計算

相關情境分類與計量：天文意象 4/8

天文景觀分為日、月、星、銀河四種項目。

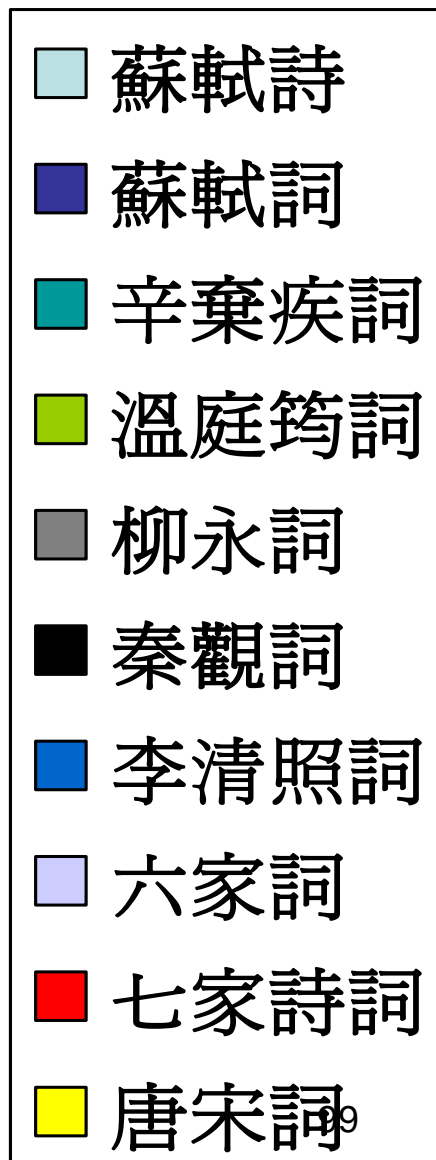
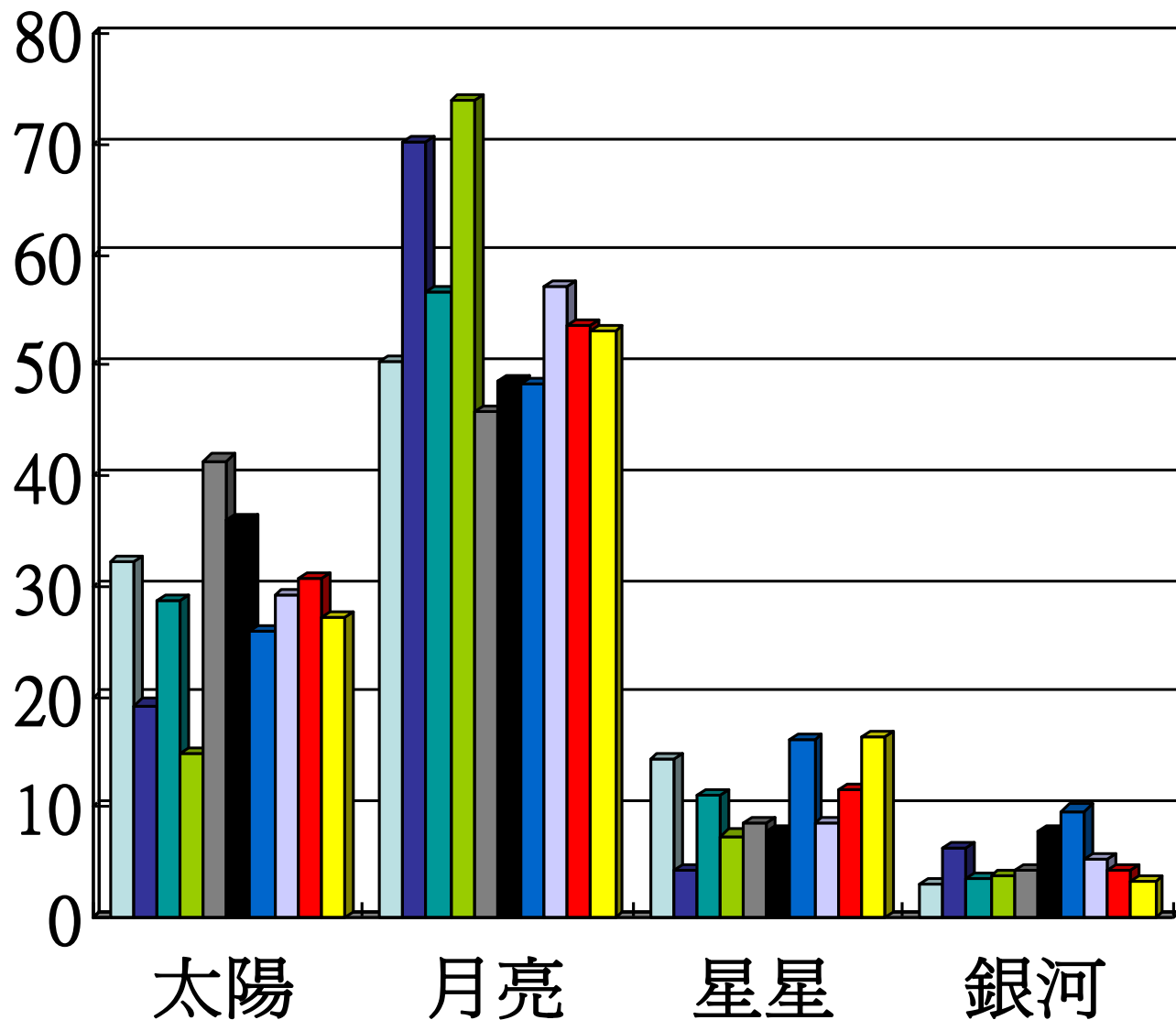
一、月亮

- 從統計各家詩詞使用的次數及比例可知，使用頻率最高的都是月亮，其次是日。

二、星與銀河

- 蘇軾詞寫銀河多於星。
- 秦觀詞銀河與星一樣多。
- 其餘都是星多於銀河。

附表八：日、月、星、銀河等天文景觀比例



詩詞文學的情感分類及情感計算

相關情境分類與計量：天文意象 5/8

- 詩人曾將月與花、人、酒並列為人世間美好的事物，因此而有「縱使有花兼有月，可堪無酒更無人」（李商隱〈春日寄懷〉）、「花間一壺酒，獨酌無相親。舉杯邀明月，對影成三人。」（李白〈月下獨酌〉）等作品出現。
- 但是月亮所引發的情感，多數是「海上生明月，天涯共此時。情人怨遙夜，竟夕起相思。」（張九齡〈望月懷遠〉）的二地相思之苦。
- 或是「此生此夜不長好，明月明年何處看。」（蘇軾〈陽關詞三首：中秋月〉）的人事無常之感。
- 即便是也有「唯願當歌對酒時，月光長照金樽裡。」（李白〈把酒問月〉）的灑脫，畢竟不多。
- 整體而言，月亮所引發的情感，仍是負面多於正面。

詩詞文學的情感分類及情感計算

相關情境分類與計量：天文意象 6/8

- 詩人筆下所寫的月亮，多數使用「月」或「明月」、「月明」、「月色」等詞彙。
- 但是，七家詩詞之中，蘇軾341句寫及月亮的詩裡，常使用「殘月」（9次）、「缺月」（6次）、「孤月」（3次）、「寒月」（1次）、「冷月」（1次）等詞彙，蘇軾詞三次寫缺月，這些都是負面悲傷詞彙。
- 柳永詞中也是七次寫殘月，一次寫冷月。溫庭筠二次寫殘月、辛棄疾殘月二次、缺月一次，李清照一次、秦觀寫殘月、月寒、月冷各一次。
- 柳永等人寫殘月，或因自身的異鄉孤寂，或寫情感上的二地相思。

詩詞文學的情感分類及情感計算

相關情境分類與計量： 天文意象 7/8

- 蘇軾62歲三度被貶，初到儋州第一年歲末，心境淒愴，念及子由，寫〈十二月十七日夜坐達曉，寄子由〉詩：「燈燼不挑垂暗蕊，爐灰重撥尚餘薰。清風欲發鴉翻樹，缺月初升犬吠雲。閉眼此心新活計，隨身孤影舊知聞。雷州別駕應危坐，跨海清光與子分。」從詩題可知當時是農曆十七，月仍當圓，但蘇軾心中、筆下的月卻是「缺月」。
- 65歲獲赦北返，寫〈六月二十日夜渡海〉詩：「參橫斗轉欲三更，苦雨終風也解晴。雲散月明誰點綴，天容海色本澄清。空餘魯叟乘桴意，粗識軒轅奏樂聲。九死南荒吾不恨，茲游奇絕冠平生。」農曆二十的月色，應已清輝漸減，但此時的蘇軾，渡海北返，心境已如烏雲散盡，眼中所見便成了「雲散月明」一片清朗的景象。
- 可見詩人筆下、眼前、心中的缺月、殘月，未必等同於天上的缺月、殘月。

詩詞文學的情感分類及情感計算

相關情境分類與計量：天文意象 8/8

- 七家詩詞寫及「星」的比例僅次於「月」。
- 除了泛寫「星」之外，以星座之專名書寫者，頻率較高的是牛郎織女星、金星系列的太白長庚星以及北斗星。
- 牛郎織女星所引發的情懷，當然是有情人相隔二地的思念之苦。

詩詞文學的情感分類及情感計算

相關情境分類與計量：動植物意象 1/4

- 以「鳥獸蟲魚草木」等動植物做為詩歌比興的技巧，是《詩經》以來詩歌創作的傳統。
- 有些動植物在詩人反覆描寫之後，已經有了固定的意象，如柳樹的離別意象，麋鹿的隱逸意象，松竹的高潔意象等等。
- 詩人們最常寫到那些動植物與其作品的內容及風格有密切關係，也與作者自身的性格及遭遇有關。

詩詞文學的情感分類及情感計算

相關情境分類與計量：動植物意象 2/4

- 從附表一的分析及統計，可以看出詩及豪放詞描寫最多的動物是馬。
- 婉約派詞人偏向描寫體型比較小的動物，尤以黃鶯、燕、雁、鴛鴦居多。
- 植物方面，七家詩詞寫得最多的是柳、菊、梅。

詩詞文學的情感分類及情感計算

相關情境分類與計量：動植物意象 3/4

- 柳樹具有離別的意象，古人送行時折柳相贈，可見於《三輔黃圖》的記載：「霸橋在長安東，跨水作橋，漢人送客至此橋，折柳贈別。」
- 漢樂府詩題有〈折楊柳〉，是漢橫吹曲所留下的十曲之一。
- 魏、晉、南朝宋梁陳有〈折楊柳〉十首，〈折楊柳行〉四首，唐代有五十三首，寫的都是離別之苦。
- 唐宋詞以兩性情感為主軸，在柳永之前，詞人筆下的女性多數是歡場女子。在這些逢場作戲，送往迎來的歌樓妓館之間，柳樹被大量的描寫，用來作為離別的象徵。

詩詞文學的情感分類及情感計算

相關情境分類與計量：動植物意象 4/4

- 郭茂倩《樂府詩集》5,156首，寫到柳樹的句子有256句，佔4.97%。
- 全唐詩51,170首，是唐宋詞的二倍多，寫到柳樹的句子有3,055句，佔5.97%。
- 唐宋詞總計24,226闕，寫到柳樹的句子有3,144句，佔12.98%。
- 宋名家詩網站所收43,943首宋詩，寫到柳樹的句子有1,778句，佔4.05%。
- 唐宋詞寫柳樹的頻率比唐宋詩高出很多，柳樹同時也是唐宋詞出現頻率最高的植物，與唐宋詞主題描寫兩性負面情感偏多的現象相符。

作者/類別	季節/%	天候/%	氣溫/%	晝夜/%	天文/%	動物	植物
蘇軾詩 33.71/66.29	春天60.39	陰雨89.43	冷寒72.97	夜晚66.42	月亮50.22	馬鶴牛	竹松柳
蘇軾詞 25.07/74.93	春天71.96	陰雨76.83	冷寒77.14	夜晚82.73	月亮70.19	馬燕鶯	柳梅楊
辛棄疾詞 27.84/72.16	春天73.45	陰雨86.73	冷寒81.25	夜晚81.25	月亮56.64	馬燕雁	梅柳菊
溫庭筠詞 8.57/91.43	春天87.50	陰雨95.24	冷寒75.00	夜晚70.59	月亮74.07	鶯燕雁	柳杏梅
柳永詞 31.90/68.10	春天53.52	陰雨61.90	冷寒55.00	夜晚76.06	月亮45.69	馬鴛鴦鶯	柳桃楊
秦觀詞 19.74/80.26	春天70.59	陰雨76.74	冷寒82.35	夜晚79.55	月亮48.54	燕鶯雁	柳梅桃
李清照詞 10.29/89.70	春天80.00	陰雨85.00	冷寒93.33	夜晚75.76	月亮48.39	雁鴻馬	梅桐柳
六家詞 26.11/73.89	春天70.70	陰雨77.96	冷寒75.27	夜晚79.89	月亮57.08		
七家詩詞 30.44/69.56	春天66.31	陰雨83.94	冷寒73.95	夜晚73.29	月亮53.61		
唐宋詞 26.99/73.01	春天73.72	陰雨75.26	冷寒72.97	夜晚70.25	月亮53.19		

情感表達的媒介與方式 1/3

- 由上一個分析統計可知，在以描寫負面情感為主軸的唐宋詞，影響其感受的季節、天候、溫度、晝夜時間等外在環境條件是一樣的。
- 用來寄託其心中情感的天文、動物、植物等外物元素也是相近的。
- 天文物象的月亮，集中於表達思鄉與相思之苦，因此各家作者都一樣。
- 用以寄託情感的動物，則豪放派詞人與婉約派詞人有所不同。
- 離別是引發詩詞負面情懷的主因，因此柳樹是多數詩人詞人用以寄託情感的植物。

「凍雲黯淡天氣，扁舟一萬怒更舉。舟渡。處起帆南，一殘敗岸。遊相閣後懷凝斷。」（柳永〈夜半樂〉）

凍葉壑濤聞泛望村漁落兩行到浪竟歲杳長樂）

雲，千漸商畫中，人，三客此萍何晚杳天）

黯，巖息旅鷁酒數鳴衰三、因難據歸神暮）

淡，相、旆行榔楊，含念駐。期京。」

天離越樵呼。翩閃霜歸掩浣羞，慘阻路。」

氣江溪風。翩翩樹去映沙笑繡歎離。」

，渚深乍片過，。遊相閣後懷凝斷。」

扁舟渡。處起帆南，一殘敗岸。遊相閣後懷凝斷。」

一萬怒更舉。舟渡。處起帆南，一殘敗岸。遊相閣後懷凝斷。」

從這一闕詞可以看到作者的心境隨著外物的轉移而變化。

秋出了、敗、己、空懷最的歎

的「了、敗、己、空懷最的歎」而離，中感歎

興到「自駐掃慘」色而

氣乘是日、映起難一「阻暮進

天「但殘」掩想萍此從期向，

淡，「去楊，浪至，歸望」。

黯葉快到歸衰」，興下晚」路據

雲一輕看榔「女拋之直歲眼京無

凍舟境為鳴、遊輕遊轉恨淚神寧

在，扁心因人」沙閣出急空凝杳丁

永晨遊昏「零」晨情到後「杳約

柳

從這一闕詞可以看出作者用以烘托情感的元素，包括：

- 季節（深秋）
- 天氣（凍雲黯淡天氣）
- 自然景觀（敗荷零落，衰楊掩映）
- 時間（黃昏—引發愁緒的時間點）
- 植物（霜樹、敗荷、衰楊、浪萍）
- 動物（斷鴻）等相關情境

正是《文心雕龍·物色》篇所說的：「春秋代序，陰陽慘舒，物色之動，心亦搖焉。」、「情以物遷，辭以情發」的寫照。

小結

1/2

一、在以描寫二性情感的唐宋詞作品裡，寫負面情感遠多於正面情感。

二、作品中用以表達情感的因素包括季節更迭、天候氣溫、晝夜時間、景觀變化、自然天象萬物等客觀外物：

1. 季節景觀：春天 > 秋天

2. 天候變化：陰雨 > 晴天

3. 氣溫升降：冷寒 > 暖熱

3. 晝夜時間：夜晚 > 白晝

4. 天象景觀：月亮 > 星 > 銀河

5. 自然萬物：動物：馬、鶯、燕最多；植物：柳樹最多

分析出情感表達的因素，也同時得出情感表達的規則，有了規則，便能夠透過計量的數據識別情感。

小結 2/2

- 無可否認，文學裡情感的表達與人類情感的表達都是複雜多變且因人而異的。
- 分析的規則與歸納統計的數據只能代表情感的共相，無法涵蓋個別的差異。
- 如前所述，「酒、花、月」被多數人認為是人世間美好的事物，但是李煜亡國後卻無奈的發出「春花秋月何時了」的哀嘆悲鳴。
- 除了情感表達的方式多變，歸結到文學作品本身，除還有語意、語用、語境的問題以及語法的變化，在在都是影響本研究的因素。
- 電腦不能解決所有的問題，本研究嘗試以情感計算及語意概念的建構詩詞情感表達與識別研究，從影響情感後續更多的研究。

第五部分

遭遇的困難與未來的發展

遭遇的困難與未來的發展 1/7

詩詞語言的特殊性增加了詩詞詞彙切分及語意標記、語意分類的困難。

詩詞語意概念的分類，除了必須符合知識認知層次的語言知識分類，還需顧及情感認知層次的意象取向分類。（柳：知識認知＝植物；情感認知＝離別）

詩詞的詞彙以雙字詞居多，但是詩詞句子的字數多數是奇數，所以多數奇數字數的句子之中都會有一個單字詞，單字詞的詞義比雙字詞更豐富多變，這也增加了詞義標記的複雜性。

詩詞裡典故用語的標記方式比一般用語困難。

以某些詞彙共現以烘托某種情境的情境用語，其詞義無法在單一詞彙之詞義顯示或涵蓋，無法在單一詞彙單獨標註詞義，哲理詩尤其困難。

這些特殊詞彙的語意，需要仰賴更多相關的研究，用更多分析、統計、歸納的方法尋找規則。

詩詞語意標記的困難，還包括詩詞喜用多義字來創造聯想的空間，以引出無窮的趣味及意在言外的情境。

宋代以後，受到理學發達及禪宗思想的影響，有很多的哲理詩，需仰賴讀者的哲理思維領悟，當然也不是單獨的詞義標記所能夠清楚描述的，想要借助詞義標記或語意概念分類讓電腦能理解，其難度又比深層義及言外義更高，在分類上，只能以詩的主題作分類。

遭遇的困難與未來的發展 4/7

蘇軾〈題沈君琴〉：「若言琴上有琴聲，放在匣中何不鳴？若言聲在指頭上，何不於君指上聽？」

楊萬里〈戲筆〉：「野菊荒苔各鑄錢，金黃銅綠兩爭妍。天公支與窮詩客，只買清愁不買田。」

這些哲理詩正是劉大魁所說的「只可意會而不可言傳」的作品。

從以上所述可知，可以透過語意標記及語意概念分類，讓電腦從分辨字形提升到理解詞義，使電腦的認知能力更接近人的知識體系。

透過情感特徵的分析及分類，使電腦具備情感識別及表達的能力。

人的知識體系分為「對知識的認知」、「對情意的感知」、「對哲理的辨知」，認知是表述的層次，感知需要聯想輔助，哲理詩的解讀體會，卻不是單純的邏輯思維能解決，也就難以用語意標記處理。

本研究從語意學及使用者的認知觀點，以中國古典詩詞網站的設計作為實證的研究，初步建立中國古典詩詞語意標記資料庫及中國古典詩詞語意概念分類表。

期待後續的研究，能進一步彙整、分析、歸納中國古典詩詞語意類別，再與共現詞彙比對，期望能從中建立中國古典詩詞主題分類的方法與規則，以彌補《唐詩分類大辭典》等前代類書分類的缺失與不足。

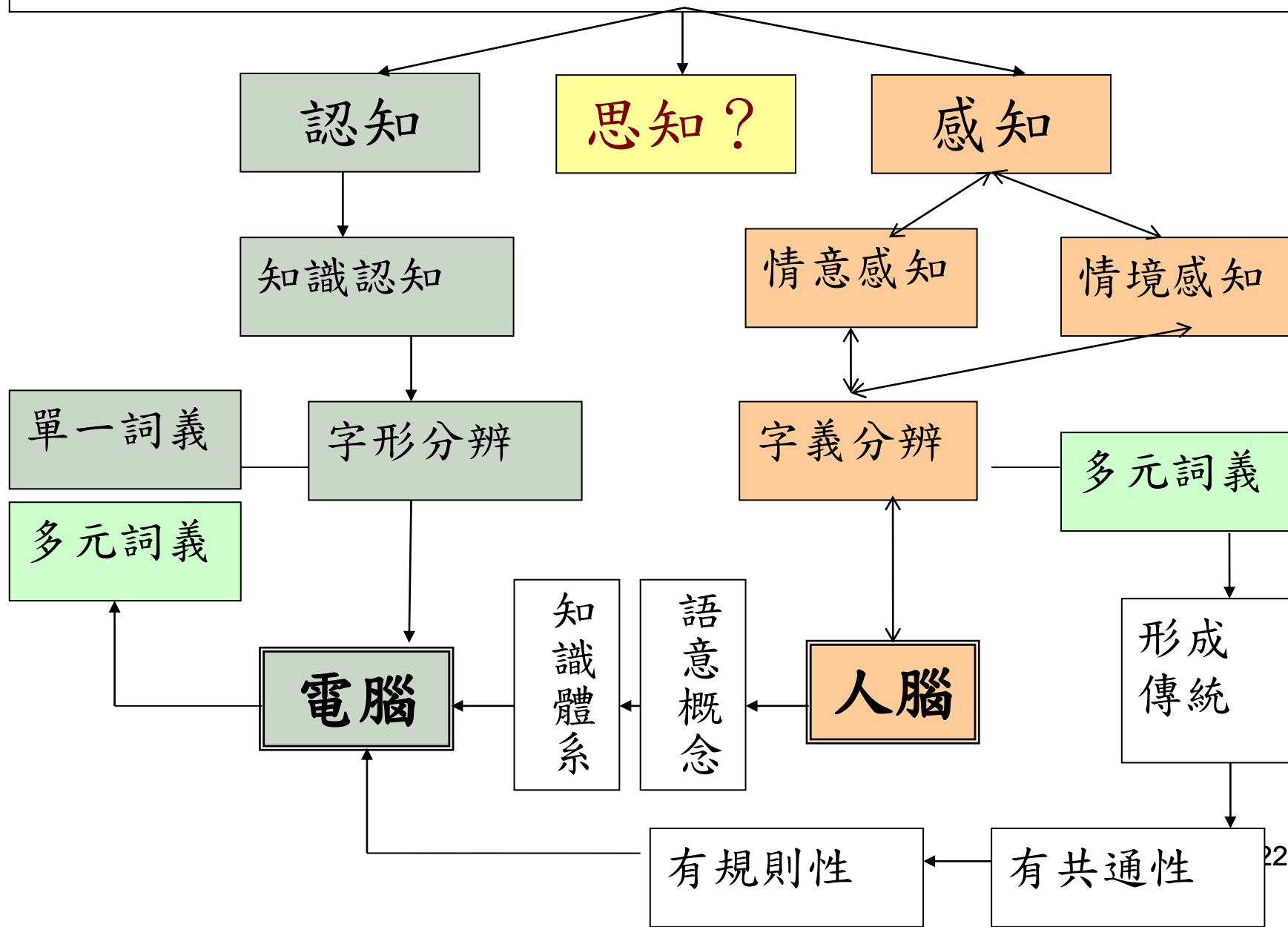
遭遇的困難與未來的發展 7/7

本研究目前只完成上下階層的樹狀結構，尚未處理橫向階層的相關性結構，這些都有待後續的研究補足。

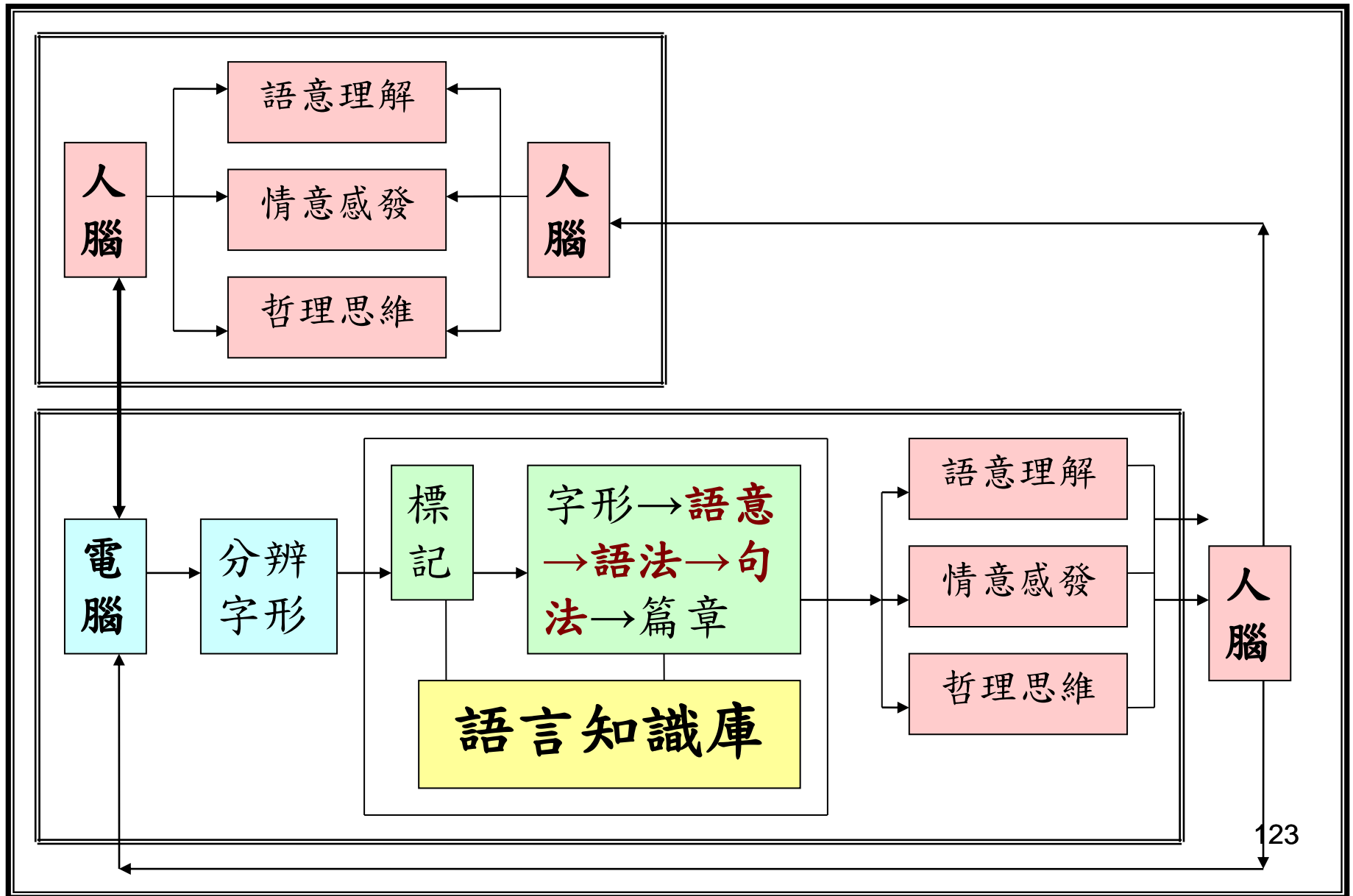
語意標記部分，也還需仰賴語／句法標記、詞性標記來提升理解詩詞之詩意與情境的能力。

建構一套從單字及詞彙的語意、到語法、語用、語境的詩詞語言知識庫，這是一條漫漫長路，仰賴更多學科的整合，也仰賴更多研究的補強。

如何讓電腦更接近人腦：以詩詞賞析為例



未來的發展：建立語言知識庫



感謝聆聽 敬請批評指教

講義網址：

網路展書讀：<http://cls.hs.yzu.edu.tw/>