

判讀報告本體論知識系統建構之研究 ——以甲狀腺疾病影像為例

蘇國璋

高雄第一科技大學資訊管理系

許蒞彥

高雄第一科技大學資訊管理系

林惠娥

長庚紀念醫院醫療事務處

摘要

目前在各醫療院所最常見的資料搜尋方式係採用關鍵詞搜尋法，此法最大的缺點之一在於資料儲存前未依病灶特徵做分類，以致產生搜尋結果資料過多或遺漏的問題。本研究為應用本體論的技術特點，以台灣南部某醫學中心的甲狀腺疾病影像判讀報告為例，分析其知識內容，將隱藏於報告中的專家知識，依疾病特徵及器官之間相互的關聯性加以分類，並標示腫瘤或器官等各類別的不同的屬性，建立判讀報告知識庫，透過它來描述複雜的知識及其間相互的關聯；並利用語意分析的方式將知識庫建置於 Stanford 大學 Medical Informatics Center 提出的 Protégé 知識平台上。本研究之知識內容層次的分析為著重在標準化及系統間的語意溝通，由使用者描述抽象問句，透過推論機制系統即可由知識架構中，有效找出符合語意的文件。

關鍵字：本體論、知識管理、語意網絡



A Study in Developing an Ontology-based Knowledge System for Thyroid Gland Image Reports

Kuo-Wei Su

Department of Information Management, National Kaohsiung First University of Science
and Technology

Li-Yen Shue

Department of Information Management, National Kaohsiung First University of Science
and Technology

Hui-O Lin

Department of Medical Administration, Kaohsiung Chang Gung Memorial Hospital

Abstract

Most medical information systems in this country use keywords to search for relevant information. The major issue with this approach is the fact the keyword matching does not take into consideration the intended semantic of users, which may include elements of disease's symptoms, signs, and others. This study applies the ontology methodology to analyze image reports for thyroid disease; according to its symptoms, tumor classifications, relations with other syndromes, and create a knowledge structure that interlink various disease elements in the database to facilitate relevant information retrieval. We, based on the actual cases from a major hospital, develop the knowledge taxonomy by analyzing the report contents starting with major disease symptoms and related body organs. The inter-relationships between symptoms and body organs were developed through attributes of various taxonomic classes. In building the system, we apply semantic network to analyze and develop relationships between knowledge items, and utilize the knowledge platform Protégé, developed by the Medical Informatics Center of Stanford University, for user to describe the abstract question. This can find out the file which accords with the user's purpose effectively from the knowledge structure through the inference mechanism.

Key words: Ontology, Knowledge Management, Semantic Network

壹、緒論

資訊科技的發達是促成醫療院所邁向知識型組織的動力，而知識管理的過程在於激勵組織成員創新智慧與經驗，利用資訊科技做有效、即時的紀錄，編碼、分類與儲存，並擴散給需要這些知識的成員；同時這些有價值的知識還必需不斷地更新，透過管理的機制，保存知識的延續性，使繼承的知識對組織更有價值，以成為醫院的重要資源（王德玲譯 民 92）。一般醫務人員或臨床研究人員，在大量報告的環境下，欲於短時間內找到某一段期間內同時具備相同疾病特徵的判讀報告並不容易，且就影像醫學而言，某一腫瘤壓迫到其他器官或影響周邊器官時，同時可能有多個器官功能受到不同程度的影響，而這些影響程度及範圍並非局限於單一部位，為解決報告搜尋上的困境，資訊人員常將報告電子化，並提供資訊檢索方式，讓使用者能在大量的報告中搜尋，但提供過多的報告或因判讀報告中醫師在報告內容所使用的語意表達方式不同而有所遺漏，其原因之一是搜尋引擎並不能真正瞭解使用者想要的報告為何，以致於搜尋結果還須要花費大量的時間進行篩選。

為了解決大量判讀報告所衍生的困擾，大部份的醫療院所均會導入檢查報告系統，它是以文件為基礎（document-based），利用全文檢索方式以關鍵字比對搜尋，其功能可分為報告類別編碼、檢查項目編碼、判讀內容對應編碼、報告儲存以及報告搜尋等五個步驟；此方式最大的問題在於「無法瞭解使用者的語意」，使搜尋結果資料量過大或非全為使用者所需，或因索引用語不同而造成資料遺漏，導致使用者必須再次篩選。針對判讀報告上疾病特徵的搜尋，大多數醫療院除以關鍵字搜尋外，另一種資料採礦（Data Mining）方式係以 ICD-9-CM 編碼來搜尋，內容包括疾病代碼列表說明及手術與處置代碼列表說明，較適用於表達病患在某一種療程或處置方式的狀況，但無法由疾病分類的方式完全涵蓋，對於搜尋影像特徵判讀報告資料，並非全然適用（White, Brickner, & Scannell 2004）。

在報告管理的機制下，若能應用語意式網絡（Semantic Network）的技術可以解決報告過多或遺漏的問題；當類別分的愈細，並且考慮判讀報告醫師對於報告內容寫法上之「用法不同但語意相同」的問題，愈可解決報告搜尋的問題（Jacquelinet et al. 2003）。透過語意式網絡的技術，資訊系統將可瞭解報告內容的語意，並利用語意式搜尋（Semantic Searching），讓系統找到使用者真正想要的文件，這將縮短整個搜尋及篩選的過程。若要使系統自動分析與搜尋使用者所需的資料，資訊人員必需將人類的想法轉換成系統所能瞭解的模式，就是先將專業領域知識（domain knowledge）告訴知識庫系統，再由系統分析每份判讀報告的知識，系統即可瞭解並管理報告。而本體論（Ontology）即指透過語意的表達來陳述個體與個體相互之間的關聯，同時亦可透過語意來聯結彼此間在同一知識平台上運作的法則，為一個概念需求上可彼此共享的理論（Gennari et al. 2003）。

其次，在判讀報告的記錄中，腫瘤(tumor)大小的記錄可說是相當重要的一部份，因為腫瘤大小關係到患者的治療方式與給藥方式，其準確度的要求必需是 100%，不容許有絲毫誤差。報告中常會出現與長度單位有關的資料，而這些長度計量單位會隨著判讀報告的醫師習慣不同而出現不同的長度單位；在報告記錄的管理上，使用者以 keyword 搜尋即可能因為無法同時透過不同的長度計量單位進行搜尋，而找不到想要的報告記錄。另一方面，判讀報告的記錄上也常出現諸如 tumor >0.5 cm 或 0.4 x 0.5 x 0.6 cm tumor 等描述腫瘤在三度空間立體定位的影像中所顯現一維、二維甚至三維的報告記錄，對於大於某一長度或者對腫瘤大小的三度空間搜尋等，均為目前資料庫搜尋上無法作到，亟待突破之處。

本研究經初探南部某醫學中心及文獻探討發現目前甲狀腺疾病判讀報告系統管理的四大問題：(1)資料庫儲存方式不佳、(2)無法作到特徵的全面擷取、(3)描述疾病之專有名稱有待統一、(4)長度計量單位整合問題；針對這四大問題，本研究以南部某醫學中心為例，提出甲狀腺疾病判讀報告知識庫的架構，應用本體論的技術，分析其知識內容，並利用語意分析將判讀報告知識庫建置於 Stanford 大學 Medical Informatics Center 提出的 Protégé 2000 知識平台上。研究目的摘要說明如下：

- (1) 分析判讀報告文件的特性，依疾病之影像特徵分類儲存，以組織甲狀腺疾病在影像特徵的知識內容之間的關聯性，並將其建置於 Protégé 2000 之中，來分析甲狀腺疾病影像判讀報告知識的內容；
- (2) 針對單位整合問題，本研究運用 Protégé 2000 的特性，在系統內設立不同的資料表格加以克服，超越傳統資料庫所無法作到的長度計量單位整合；
- (3) 提出一個能找出腫瘤大小三維空間的知識庫系統，以掌握腫瘤大小與被侵犯的週邊組織之間的關聯性；
- (4) 利用 Ontology 的技術，建置一個能瞭解語意的甲狀腺疾病影像判讀報告搜尋知識系統，以提高研究人員對疾病特徵的搜尋比率而縮短研究資料的搜尋時間。

研究方法首先劃分判讀報告內容綱要（病患基本資料、檢查目的及病程摘要、檢查項目或疾病分類以及檢查判讀內容等四大類別），爾後以所劃分的四大類別為主軸，分別進行甲狀腺疾病判讀報告知識分析其屬性及其關連性，接著應用 Ontology 技術來完成甲狀腺疾病判讀報告知識庫系統的整合，最後針對已建置完成的甲狀腺疾病判讀報告知識庫，以使用者需要的查詢進行推論，透過邏輯表示法將判讀報告文件中會遇到的問題轉換成系統能了解的表示法，由資訊系統進行推論，進而找出使用者想要的病患資料，此為系統評估的階段。



貳、文獻探討

一、甲狀腺疾病

甲狀腺是頸部前面的一個很重要的內分泌腺體，它的形狀就像一隻蝴蝶，蝴蝶的兩邊翅膀就是甲狀腺的左右兩葉，蝴蝶的身體就是甲狀腺兩葉的峽部，有些人在峽部尚有突出的錐葉，彷彿是蝴蝶的觸鬚，兩葉的門緣各有一對副甲狀腺，就像是蝴蝶翼上的斑點。甲狀腺的峽部正好在環狀軟骨下方，而兩葉則貼於甲狀腺軟骨兩側的下部。內分泌腺又叫無管腺，乃是指一些腺體，它的內分泌物直接分泌到血管，藉著血液循環，而作用於身體的其他部位。人的內分泌腺體除了甲狀腺外，尚有腦下垂體、副甲狀腺、胰臟、腎上腺、卵巢和睪丸等。內分泌腺的分泌物叫荷爾蒙，或叫激素。甲狀腺荷爾蒙是維持我們身體健康所必須的，它可以促進體內各種組織的新陳代謝（林素蓉譯 2000）。

此外，甲狀腺疾病在臨床診斷上常藉由下列影像之照攝來協助診斷（楊宏智、林宏達 1995）：

1. 胸部 X 光後前照及側照（Chest PA & LAT），能夠顯出氣管移位、氣管被壓迫、鈣化或胸骨後腫瘤。
2. 超音波（Ultrasound, ECHO）對腫瘤的影像檢查，對於單一或多發，囊狀或實質的腫瘤，以及配合細針抽吸切片做細胞學檢查，對於偵測結節是否為良性、懷疑或惡性等非常有幫助。
3. 電腦斷層（Computer Tomogram, CT）或核磁共振影像（Magnetic Resource Image, MRI）對於大且侵襲性、胸下的甲狀腺腫，尤其當腫瘤無法被適當地評估時，以及配合肺功能檢查或氣管氣體交流檢查，對位於胸骨後的巨大甲狀腺腫非常有幫助。

二、醫療資訊系統

區域醫院及教學醫院通常將醫療資訊系統（Hospital Information System, HIS）劃分為七大部份，分別為 1.醫院管理資訊系統（HIS）、2.檢驗管理資訊系統（LIS）、3.臨床資訊系統（CIS）、4.醫療影像系統（PACS）、5.行政管理系統（MIS）、6.放射線科系統（RIS）、7.患者服務系統（FRM）等七大系統。其中，RIS 包含了放射線科管理、檢查排程系統、檢查報告系統等三大部分；放射科管理的工作有產生影像、分析影像、管理影像及其他相關資訊處理；檢查排程系統主要功能為藉由排程的登錄及未到診原因的分析，以提高服務品質及效率（吳昭新、李友惠等 2002）。以下為針對與本文有關的檢查報告系統，PACS 與電子病歷（Computer-based Patient Record, CPR）做詳細說明。

(一) 檢查報告系統

檢查報告系統有編碼輸入、關鍵字、摘要、連結教學檔、檢查項目編號輸入、報告查詢、檢驗查詢、來源別、醫師編碼以及影像連結等功能。檢查報告系統醫囑連結的重點在四個構面上，如下說明（經濟部標準檢驗局 1999）：

1. Initial Findings：病患就診時最初發現的病灶，代表獨立療程或處置開始。
2. Indications：針對病灶，臨床醫師依其經驗判斷所決定的處置方式。
3. Special Procedure：針對病灶及病患既有的疾病，依照醫囑執行處置過程。
4. Performance by：配合病情需求的執行方式。

這四個構面對判讀報告內容而言，是不可或缺的標準。此系統將顯示病患的過去與現在、手術前與手術後、治療前與治療後、病情程度的表現、療程分級的標準等資訊；同時醫囑在開立的過程中共有四項主要輸入項目，分別為：Subject（病患主訴過去病程內容）、Object（病患理學檢查結果）、Assessment（臨床醫師診斷輸入）、Indication（臨床醫師安排檢查目的及病程摘要）等四項（Bates et al. 1995）。

檢查報告系統的優點除了提供全院一致性的品質系統評審基礎外，並可藉此找出具有可輔助臨床診斷與教學研究的關鍵性知識與報告，使醫院保有其優勢及競爭力，亦可作為在院際間資料交換時之文件基礎（吳昭新、李友惠等 2002）。

(二) 影像傳輸系統（PACS）

PACS 基本架構包括影像擷取設備、影像資料庫檔案存取系統、影像資料庫管理系統、影像閱覽軟體、影像顯示工作站；可有效處理每天大量產生的醫學影像之儲存管理與查詢，協助醫師更有效的掌握各種相關的醫療訊息、縮短會診與相互諮詢的時間，提升醫療的品質（范碧玉 民 89）。對於整合醫療資訊，標準化是最重要的環節，而在影像格式的標準化包含二個部份，一是影像檔案的標準化，另一個是影像傳輸界面的標準化。在美國 ACR/NEMA 歷經多年的改進後提出 DICOM3.0 標準（Ridley 1995）；國內則於 2000 年 11 月 30 日加入「國際 DICOM 組織」（Digital Imaging & Communications in Medicine, DICOM），同時於 2001 年 6 月 22 日成立「台灣健康資訊交換第七層協定協會（HL7）」，用來傳遞醫療影像相關資訊及交換共通格式。

PACS 是影像產生、傳輸及存取的規範，Health Level 7 (HL7)則是放射線科系統（RIS）與醫療資訊系統（HIS）的規範，PACS 與 RIS 整合之溝通系由「雙向的介面」傳輸，來協助影像存取，以提高醫師的效率、減少病患醫療成本與時間；並使用 DICOM 標準規範的「傳輸原則」以及特殊程式語言，使系統的連接性更具一致性（何美玉 1996；范振添 民 91；吳昭新、李友惠等 2002）；兩者間的整合，將使 PACS 更具有智慧，知道在何時（When）該如何（How）把什麼樣的資料（What）送到哪裡去（Where），即不受時空距離的限制。但 PACS 系統與判讀報告間仍存在某些未合乎期望之處，茲簡述如下：

1. 與檢查有關的病患資料，如病人的症狀、主訴及來作檢查的原因並不完整，或是根本沒有資料可供放射線科醫師作鑑別診斷的參考。

2. 要瀏覽一個新病歷時，因影像數目太多而無法即時看到。
3. 顯示器的視野不如傳統的片架寬闊，使得比較不同系列的檢查，變得較困難且較不自在。
4. 商用的 RIS/PACS 極為昂貴，並不是每家醫院都願意負擔這個費用，而採取觀望的態度。
5. 就技術層面而言，有些廠商或檢查儀器，至今未能完全合乎 DICOM3.0 的規範，以及在 PACS 中基本資料中文化的困難，有些醫院的檢查儀仍無法全部更新為數位化。

以上這些都可說是今日 PACS 所面臨的挑戰。

(三) 電子病歷 (CPR)

CPR 是醫療產業以電子文件記錄病人診斷影像、生理訊號、檢驗報告、醫療處置影片及醫療書單等之病歷資料與報告，可供文字及靜、動態影像之記錄播放與接收者。藉著整合不同層級、屬性的醫療院所，提供資訊交換與交流建構共同的交換平台，使資料雙向傳遞交換的即時性、及病患醫療的連續性。CPR 的推廣可達到：1.提升病歷分析研究效率；2.確保醫療品質；3.增進病情監控能力（吳昭新、李友惠等 2002）。

CPR 的優勢包括下列各項：1.節省空間。2.降低行政、人事成本。3.簡化加速流程、提升效率與正確性。4.提供完備整合資料，資源共享、避免重覆浪費。5.即時線上支援提升病人照護品質。6.查詢便利、申報更容易。7.促進教學研究。在發展 CPR 仍需持續努力之要項包括：1.法源基礎。2.認證機制。3.資料交換標準。4.醫學用語標準化。5.病歷內容格式化。7.格式統一化、資料數據化。8.不可變機制。9.安全機制、保障隱私。

三、本體論 (Ontology)

(一) 定義與功能

Ontology 係指透過語意的表達來陳述個體與個體相互之間的關聯，同時亦可透過語意來聯結彼此間在同一知識平台上運作的法則，為一個概念需求上可彼此共享的理論（經濟部標準檢驗局 1999），因此，被定義為「可明確而詳盡的描繪分享的概念」。透過共享的概念，用於知識管理中，可以讓不同領域的雙方進行溝通，即知識的表達；能共同地討論事情，甚至也可以讓人與機器或機器與機器間相互了解對方所要表達的語意。

一般而言，Ontology 提供字彙及提供儲存知識實體的功能，用來描述特定領域中的專有術語及其關係（Staab et al. 2000）。在應用方面，常用於溝通（communication）、交互運作（interoperability）及軟體工程（software engineering），溝通即是知識表達及知識分享；交互運作為本體論應用於整合性系統，以使機器自動化地交換資料或資訊，亦是知識分享；在軟體工程方面，將 Ontology 應用於系統開發之中，以達到元件重用（reusing）、提高可用性（reliability）的目的，亦是知識的重用（Uschold 1996）。由

於 Ontology 具有重複使用 (reuse) 的功能，被廣泛使用在知識管理、電子商務及電子資料交換 (梁定澎 民 93; Tamma et al. 2005)。尤其是在醫療領域，電子資料的交換促進院內本身或院外醫療從業人員就相關病症資訊之交流，也提昇知識管理的內涵，不再將知識管理定位在單純的 e-learning。因此在描述 Ontology 方法論時，須先定義其功能及認知的需求，再透過語意分析 (semantic analysis) 已存在的醫療用語。

Ontology 應用到知識管理中，提供了知識編碼、儲存、維護及搜尋等支援。在編碼方面，Ontology 利用 Frame-based 或 Logic-based 語言針對知識加以描述；在儲存及維護方面，目前許多的 Ontology 工具，皆提供知識庫的維護；在搜尋方面，Ontology 可結合推理機進行語意式查詢，能夠在適當時機瞭解使用者描述的問句，並找出符合的答案 (Bechhofer & Goble 2001; Weng et al. 2006)。

(二) 醫界發展重點與相關研究

目前 Ontology 在醫療領域發展的重點如下：第一是透過 Ontology 分析描述某物件的功能、物件轉移現象及物件之間的規則；第二是透過 Ontology 將描述某病症或某專科領域的字彙格式化 (specification)、提供標準用語 (Unified Medical Language System, UMLS)，以呈現關於某種病症的相關資訊；第三是在某種專業領域上，使 Ontology 具有自動邏輯說明的能力，經由疾病專門字彙的使用，連結至相關字彙的本體論以便於萃取相關知識，讓疾病知識地圖的表達，能透過圖解或範例的呈現更趨完整 (Gupta et al. 2003)。

在醫療領域，與 Ontology 相關的應用研究極多，在不同的研究主題上大放異彩，例如：Jacquelinet et al. (2003) 成立健康研究中心，針對血液透析病患各項檢查報告建立多階層架構；Gupta et al. (2003) 透過大量的醫療系統所提供神經科相關資料建立疾病地圖；並指出 Shortliff et al. (1981) 規劃一個癌症病患化學療法的建議系統以及 Stenves et al. (1999) 運用 Ontology 來解釋蛋白質的各種不同的屬性、蛋白質所具有各項不同的功能以及蛋白質彼此之間的相似之處，提供給醫護人員在擬定相關決策時更多的參考資料等研究應用。林郁馨 (民 90) 則運用 Ontology 於口腔超音波網路教學系統之建立與評估上。

藉由上述列舉的研究應用，可得知 Ontology 在共同分享的概念下，對特定領域可作正式而且詳盡的格式化表述，藉由這些概念上的推論到抽象模組的呈現，表達人類對真實世界的所有想法與觀念。在知識庫的設計中，「知識表達」、與「知識推理」是最核心的部份，知識表達就是將解決問題所需陳述的知識用良好的結構來表達，而知識推理則是指運用知識來解決問題，好的知識推理策略對問題解決的績效將有重大的影響 (梁定澎 民 93)。本研究利用 Ontology 理論，將甲狀腺疾病判讀報告文件內容的分析著重在標準化及系統間語意的溝通，讓系統能夠明瞭人類所思考的架構及所要表達的事項。組織 metadata，並利用 Ontology 的技術結合推論機制來達到語意式搜尋，因此，應用在報告管理上是最適合的。

四、Protégé 2000

Protégé 2000 是由美國 Stanford 大學 SMI (Stanford medical informatics) 研究室所開發出來的一套以本體論為基礎的知識庫平台，可幫助使用者在特定的領域中獲取如專家系統所能提供的客製化知識。本研究即利用 Protégé 所具有之高度延展性及相容於不同的平台的特性，及其被廣泛應用於醫療領域裡許多不同的專案中而加以採用，並藉以表達甲狀腺疾病判讀報告所隱含的複雜知識內容。Protégé 2000 的設計原理是以本體論為觀念上的主軸，提供 Frame-based 的知識輸入及擷取模式，同時使用者亦可透過 Plug-in 的方式，在系統中加入 Domain Knowledge，及在推論的過程中可茲運用的相關程式，延伸推論的知識範疇；亦允許同一個專案或性質相近的專案可拆分為若干小專案分別獨立建置系統，最後再透過 Including 的方式合併為一個系統，以擴大 Domain Knowledge 的範圍 (Gennari et al. 2003)。

Protégé 2000 的系統架構共分為 User Interface、Core Protégé 2000 及 Persistent Storage 三個層級，User Interface 主要是利用預設的介面模組，可讓使用者自行定出所想要的介面 (widget)，另外也可經由 Plug-in 的方式增加其介面；Core Protégé 2000 的層級則是整個 Protégé 管理知識庫的主系統，包括了知識階層的建立及階層間的推論；Persistent Storage 包含實體儲存媒體及 Knowledge-base Mapping 兩部份，利用 Knowledge-base Mapping 將本體論及其實例對映到實體儲存媒體上，同時也可經由 Plug-in 新增儲存媒體。

由於 Protégé 具有高度延展性及相容於不同的平台，且被廣泛應用於醫療領域裡許多不同的專案中，在此將 Protégé 的優點歸納如下 (經濟部標準檢驗局，1999；胡訓誠 民 92)：

1. 高度重用性：系統化建立的資料，可重複用於相關領域的研究上。
2. 具有高度延展性：使用者可針對醫療環境需求，擴充想要的知識模組。
3. 強健的發展環境：受到良好的領域專家及知識工作者支援，可容易的建立有效率的知識系統。
4. 客製化的檔案輸出格式：使用者可針對專案需求，輸出成想呈現的語言。
5. 整合其他應用系統：可在系統間達到語意上的互通，使用者可利用擴充的架構與其他相關的專案一層層包進來，擴大系統應用範圍。

參、甲狀腺疾病判讀報告文件

甲狀腺疾病判讀報告存在著許多疾病與體內相關組織器官之間關聯的知識，本節將針對甲狀腺疾病判讀報告進行分析，找出其中內含的知識。第一將介紹甲狀腺疾病判讀報告內含的文件層級模式，探討各知識層級間相互的關聯，第二則分析報告中所記載的知識。

一、文件層級模式 (Hierarchical Model)

從甲狀腺疾病判讀報告的功能面來看，其中最重要者為 1.病患個人資料、2.檢查目的及病程摘要、3.疾病分類資訊、4.檢查項目、5.檢查報告內容等五項，由於疾病分類資訊及檢查項目在特定疾病之下，分類項目較少，故將此二項劃為同區間內（但於系統中可單獨或組合式查詢），因此共為四大類別，經由這些類別間的關聯來找出其中的因果關係，詳細描述如圖 1。

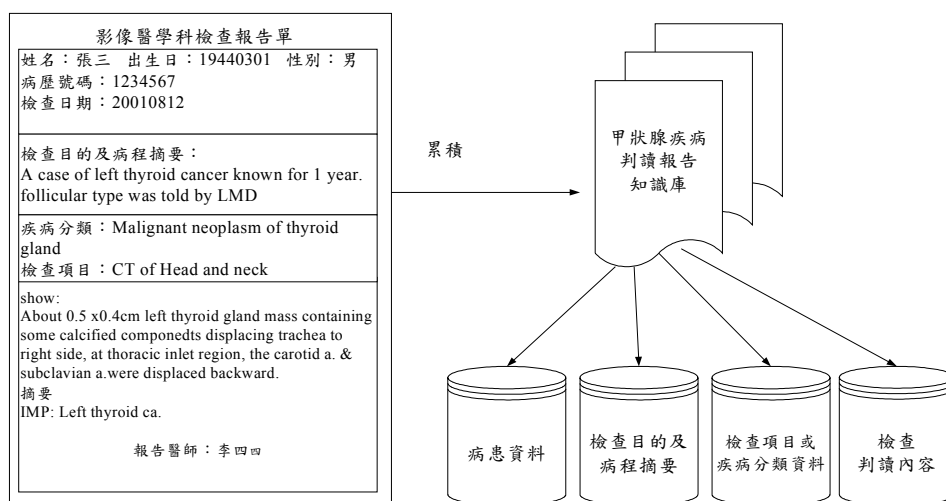


圖 1：知識庫報告內容綱要劃分

茲就上述四大類別分別說明如下：

1. 病患資料：是研究探索的中心，提供病患背景資訊，並經由病患個人資料，可再深入查詢病歷中對病患基本資料的記載或病患住院期間的護理記錄，而得知病患相關背景資訊。
2. 檢查目的及病程摘要：其功能在於它可與醫囑執行面 Initial Findings 相呼應，由病患的主述中瞭解病患就醫的目的及過去或現在的就醫史、用藥史及治療狀況。
3. 檢查項目：功能定義在醫囑執行面的 Performance by，以甲狀腺疾病為例，經由醫師問診及臨床診斷，其影像檢查項目以超音波最多，電腦斷層及磁振造影其次。
4. 疾病分類資料：針對病患在檢查目的及病程摘要中所作的陳述及臨床的診斷，透過 ICD9-CM 編碼原則來表達醫師對病患本身具有之疾病或對疾病已作之處置及臨床診斷上的臆測，換言之，即表達醫囑執行面的 Special Procedure。
5. 檢查報告內容：功能與醫囑執行面的 Initial Findings 相對應，亦為研究的重心所在。它可將檢查影像結果透過純文字的記載完整呈現，輔助評估疾病診斷之適當性、病理診斷之適當性、各種檢查之適當性及藥品使用之適當性，也是臨床研究上的主軸及醫院評鑑的重要參考指標之一。

由圖 1 知識庫報告內容綱要劃分可以明瞭，檢查目的及病程摘要、疾病分類資訊、檢查項目、檢查報告內容及病情摘要四者間存在著相互影響的關係。醫師看診過程中所記載的記錄，包括問診情形及指示之檢查...等，稱之為醫囑，其連結的重點在幾個構面上，包括(1)Initial Findings(原始發現)、(2)Indications(醫師指示)、(3)Complications(病患主述)、(4)Special Procedure(特殊處理過程)、(5)Performance by(執行狀況)等。依上述原則，首先將判讀報告知識庫定義為所有關聯的概念及系統基本知識模組的核心，它結合語意上的分析同時也支援自然語言(Natural Language)，茲將系統模組描繪如圖 2。

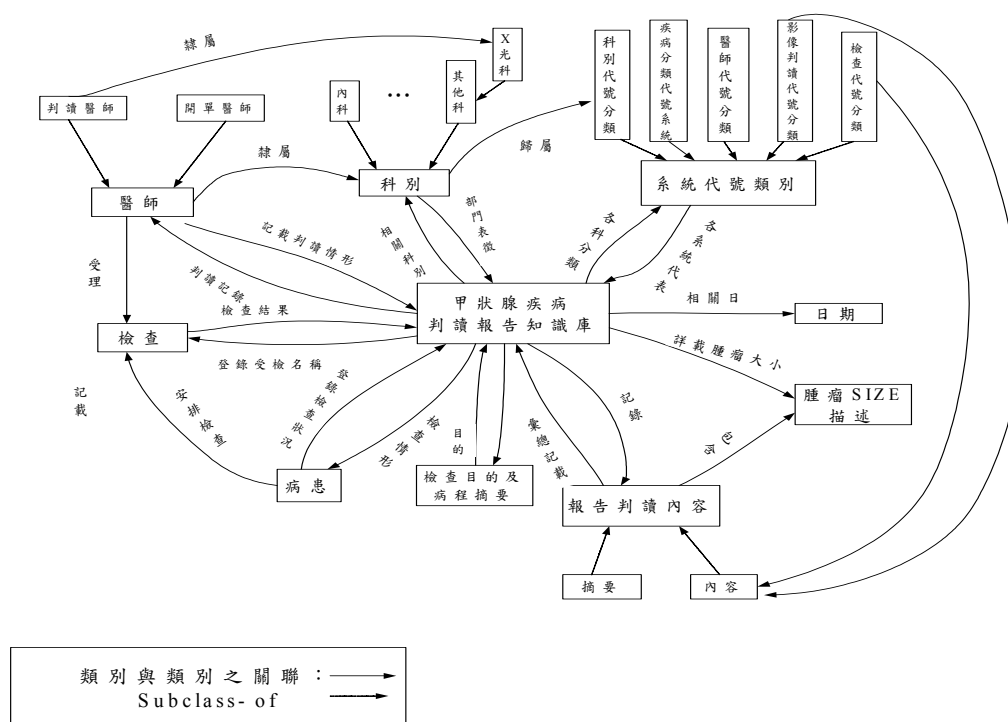


圖 2：甲狀腺疾病判讀報告知識庫類別及關聯圖

針對上述甲狀腺疾病知識庫類別及關聯圖(圖 2)，圖中與系統核心(甲狀腺疾病判讀報告知識庫)有關聯的主要類別分別說明如下：

1. 病患：每一份甲狀腺疾病判讀報告所記載的病患基本資料包括病患姓名、性別及出生日期等。
2. 檢查目的及病程摘要：報告內容常會敘述患者的就醫目的及醫師對病患就診原因的摘要說明。
3. 醫師：醫師分為開立檢查醫師與判讀報告醫師二種，任何其中一個類別的醫師均可能與研究計劃有關聯；在此先以判讀報告醫師為例，探討判讀報告醫師與研究計劃間的關聯性。判讀之主治醫師醫師必須對影像呈現之結果作一個記錄，此一記錄可能與研究計劃有關或使用研究計劃中所含括之藥品或藥材。

4. 報告判讀內容：每份甲狀腺疾病判讀報告概分為病灶的部份（Lesion）與摘要二部份，此二部份是判讀報告內容的重心，將影像結果逐一詳述，並將整份報告作摘要性記錄。
5. 報告相關日期：每份甲狀腺疾病判讀報告有相關聯日期，含醫囑開單、檢查、判讀等日期。
6. 系統代號類別：系統代號類別的次類別包括了疾病分類代號系統、科別代號分類、醫師代號分類、影像判讀代號分類、檢查代號分類等，在此以疾病分類代號系統為例，其分類資料與 ICD9-CM 國際疾病分類編碼的資料相同，並提供患者已作之處置或醫師對疾病所下的分類診斷。

甲狀腺疾病 Ontology 中的甲狀腺疾病本體論分類代號對應表關聯（圖 3），它是由圖 2 中「系統代號類別」的 Class 所延伸並加以詳細說明，分類代號對應表中的 Hierarchical Model（直線的階層式關聯），它用以描述甲狀腺疾病所有的分類代號間的關聯，包括：科別代號分類對應表、醫師代號分類對應表、檢查代號分類對應表、疾病分類代號對應表、報告判讀代號分類對應表。各分類代號對應表與系統代號類別間的關係的建立是應用 is-a（subclass of）及 has-slot（對應名稱及代號）關聯來描述，如圖 3。

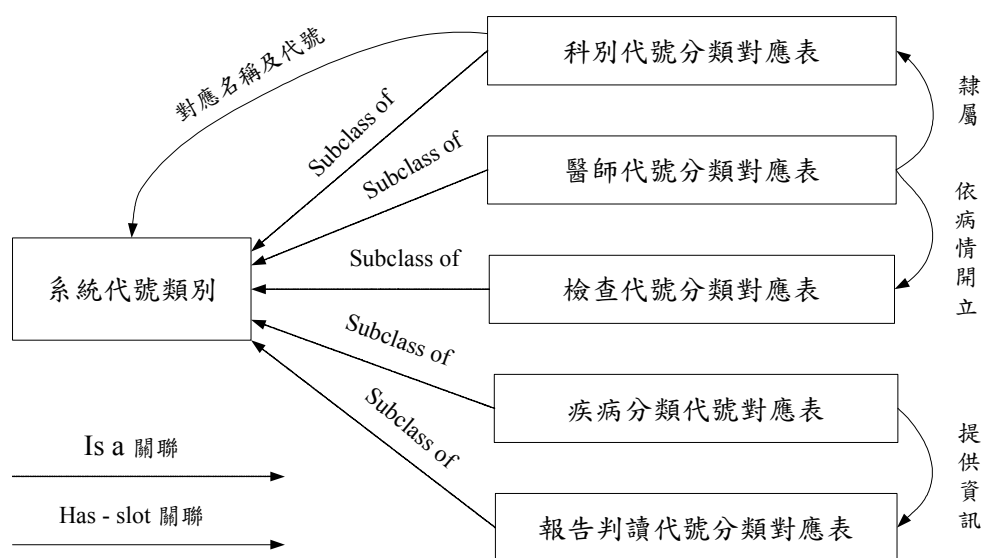


圖 3：甲狀腺疾病本體論分類代號對應表關聯

由於甲狀腺疾病影像判讀報告知識庫類別及關聯圖（圖 2）中「報告判讀內容」是知識分析的重點所在，本文亦針對此類別詳細劃分其階層架構及知識內容之間的關聯（詳如圖 4 及圖 5 所示）。

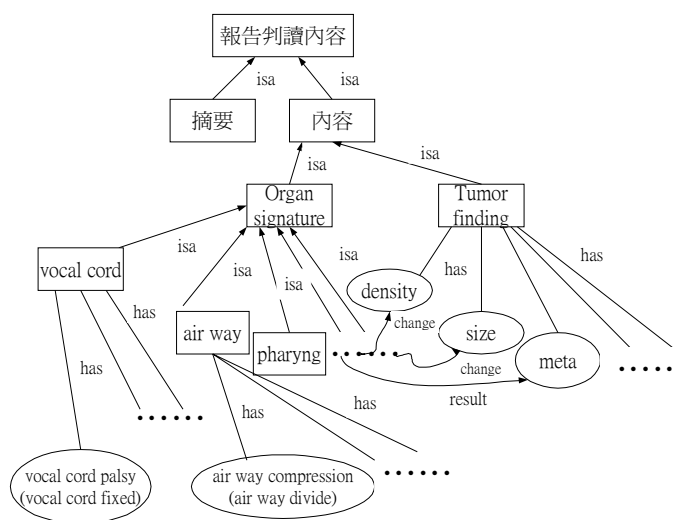


圖 4：判讀報告內容之階層架構及關聯圖

甲狀腺疾病報告判讀內容、病灶部份的內容 (Lesion)、國科會研究計劃或廠商贊助研究計劃三者之間的複雜關聯，在報告判讀內容與病灶部份的內容之間存在著一對二的特性，一份甲狀腺疾病報告判讀內容包含有「摘要」及「病灶部份的內容」二個次類別。「病灶部份的內容」與國科會計劃或廠商贊助計劃之間的關聯是透過研究計劃這個 metadata 來完成，因為研究計劃本身為一 class，所以「病灶部份的內容」須透過 metadata 的關聯來建立「病灶部份的內容 (Lesion)」與研究計劃之間的關聯，使得「病灶部份的內容」可連結到多個項目。且國科會計劃或廠商贊助計劃本身為抽象類別，並無法直接登錄與研究計劃相關的 class，所以和「病灶部份的內容」有關的項目必須為「研究計劃」的子類別，如使用材料、藥品……等，茲將「報告判讀內容」與「研究計劃」之間的關聯表示如圖 5。

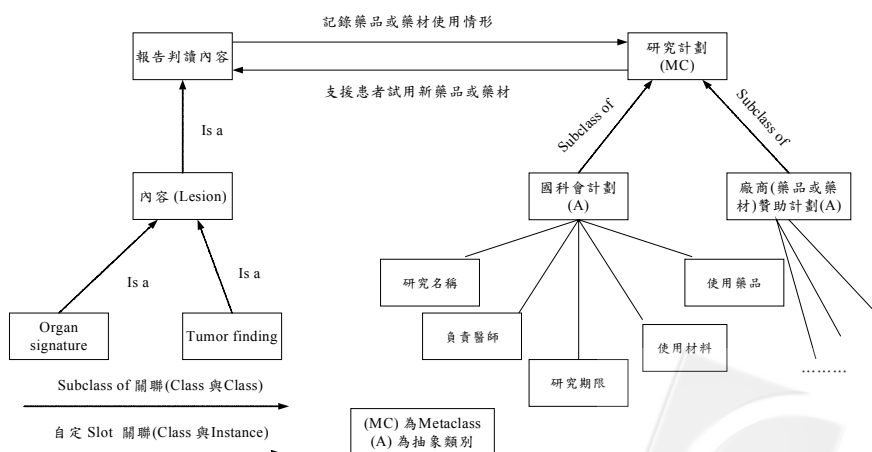


圖 5：判讀報告記載之病灶內容 (Lesion) 與研究計劃間關聯

在甲狀腺疾病影像判讀報告知識庫類別及關聯圖（圖 2）中 Hierarchical Model 表示了所有 Class 之間的階層關聯。本研究建立的甲狀腺疾病本體論知識系統中，共包括病患、檢查目的及病程摘要、醫師、報告判讀內容、日期及系統代號類別等主要 Class 及相關之 Sub Class，根據上述說明，將甲狀腺疾病判讀報告知識庫類別及關聯圖所描繪之 Hierarchical Model 以圖 6 示之。

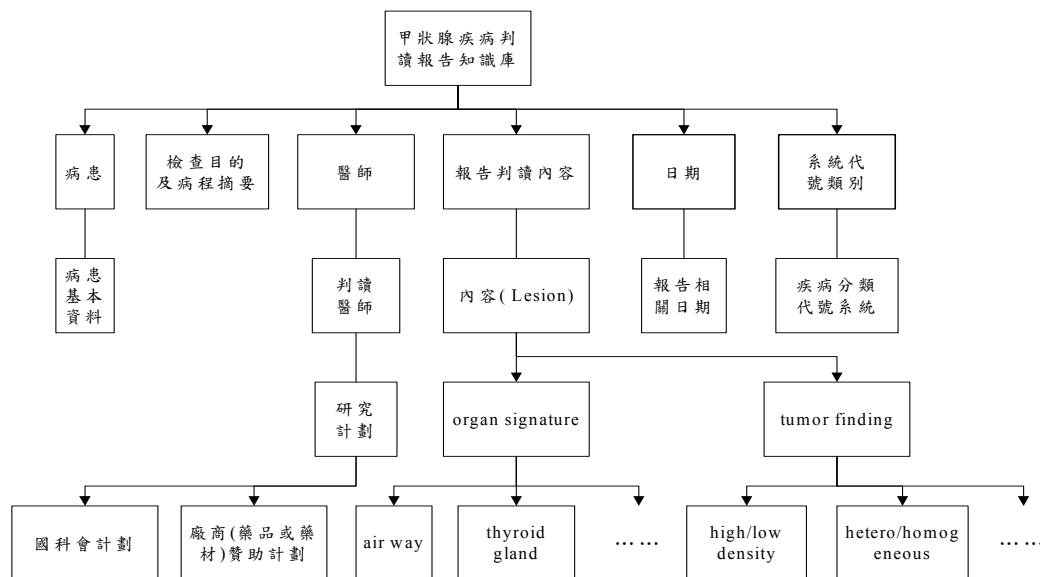


圖 6：甲狀腺疾病判讀報告之 Hierarchical Model

二、判讀報告知識分析

報告的判讀內容（圖 4）由許多片語共同組成，內容包含 1.影像特徵中的病灶（lesion）、2.摘要二大部份。影像特徵中的病灶（lesion）狀況包括對 tumor finding（腫瘤發現）的描述或相關器官特徵（organ signature）顯現的說明。摘要敘述為整篇報告內容的焦點所在，可能摘自報告內容的一小部份而與報告內容部份重複，也可能是本篇判讀報告內容的重點提示與建議方向。

從甲狀腺疾病判讀報告的知識面來看，判讀報告內容所提供許多寶貴的知識中重要者為 1.tumor finding、2.organ signature。為了使判讀報告的隱性知識能完整呈現，本研究在傳統資料累積儲存方式上加入對語意上表達的考慮，亦即系統能解讀自然語言。為了明瞭原始發現的甲狀腺腫瘤大小與其周邊組織的關聯性，就 tumor finding 方面深入探討，在單位知識上，腫瘤大小 size 包含了長(length)、寬(width)、高(height)、large、small 等，針對具體之數字描述作一維、二維、三維度空間的探索，只要判讀報告醫師所描述的腫瘤大小是相同長度，則無論以那一種長度單位的寫法記載於判讀報告的內容之中，其所代表的意義是相同的，如表 1。

表 1：判讀報告之腫瘤大小描述方式

腫瘤大小	描述方式	代表意義
Tumor size	0.5cm(5mm)	長度(一維)
	0.5cm x 0.4cm(5mm x4mm)	長度及寬度(二維)
	0.5 x 0.4 x0.3cm (5 x4 x 3mm)	長度、寬度、高度(三維)
	> 0.5 cm(>5mm) ; <0.4cm(4mm)	長度(一維)

tumor 以 Pattern (形態) 特徵來分類, 重要者包括 node、nodules、cancer...等; 以 Homogeneous/ Heterogeneous (同質性或異質性) 特徵來分類, 重要者包括 homogeneous、heterogeneous; 以 Contour (輪廓) 特徵是否很清楚來分類, 重要者包括 well define、ill define; 以 Calcification (鈣化) 形態特徵來分類, 重要者包括 calcification、cystic change; 以 meta (病灶轉移至其他器官) 的形態特徵來分類, 重要者包括轉移至 LN (淋巴)、lung (肺部)、liver (肝)……等。

檢查報告內容的分析著重於分析影像特徵中的病灶 (lesion) 與體內器官、腫瘤特徵、位置之間的關聯, 主要包括共通性、包含關聯、語意關聯、單位知識……等。共通性關聯是指用法相同語意不同、用法不同語意相同、及用法相同語意相同的相同觀念, 此關聯是判讀報告內容的共通性。利用影像判讀報告中不同的知識層級, 逐步架構起整份報告所隱含的疾病因果關係, 以系統中類別之間的相互關聯、腫瘤大小及位置的呈現、鈣化現象或發炎情形、手術前後比對、週邊器官受影響的範圍與深度情況……等影像特徵, 來掌握具有相同疾病特徵的病患彼此間所存在差異的造成原因, 可以提供使用者快速掌握影像判讀報告知識的重心, 同時對疾病特徵有更深一層的認知, 並藉此作為發展電子資料交換 (EDI) 的基礎或重點醫療的方向。

肆、系統架構

針對甲狀腺疾病判讀報告知識分析的結果, 本研究採用本體論結合推理機的技術加以解決。為貼近實務需求, 實作部份以 Protégé 為主體, 語言表述採 XML-based 為主, 強調 Schema 部份, 嘗試以語意方式應用診查報告重建符合研究需求之 Metadata 資料, 而非僅著重將診查報告轉成所要的格式 (Kim et al. 2002)。本章節將介紹本研究所提出的系統架構及相關的推論模組。首先, 介紹整體性系統架構, 並針對架構中每個模組加以介紹; 接著, 介紹 PAL (Protégé Axiom Language) 推論模組, 它整合 Axiom Evaluator 與 PAL 系統, 能使整合後的 PAL 系統如推理機般運作; 最後, 介紹在系統中利用數值型模組, 解決單位整合的問題。

一、系統架構與模組功能

本研究利用 Protégé 2000 及 PAL 推論模組來建立一甲狀腺疾病判讀報告知識系統, 用以解決甲狀腺疾病判讀報告管理四大問題: 資料庫儲存方式不佳、無法作到特

徵的全面擷取、描述疾病之專有名稱有待統一及長度計量單位整合問題。整個系統架構共分成八個模組（圖 7），茲將每一個模組說明如下（胡訓誠 民 92）：

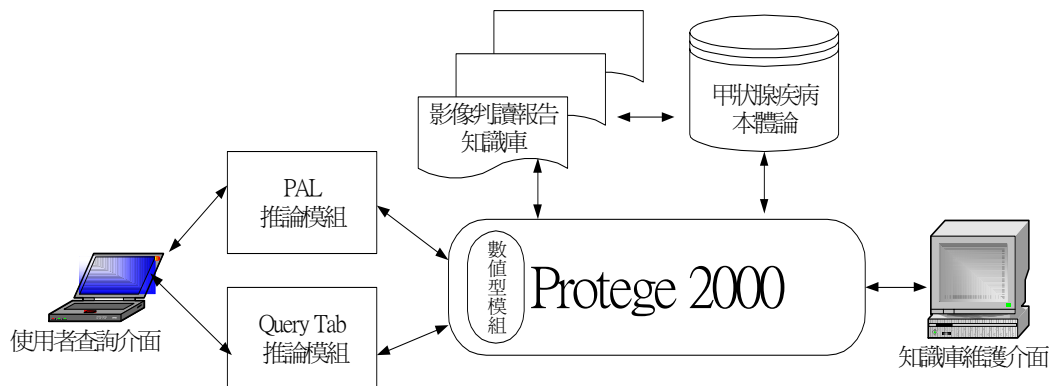


圖 7：系統架構圖

1. Protégé 2000：是本系統的核心模組，透過 Protégé 2000，甲狀腺疾病判讀報告知識庫及甲狀腺疾病本體論提供系統所需知識的分類、各類別的屬性、類別與類別間的關係以及各類甲狀腺疾病判讀報告文件之間的關係。使用者可透過知識庫維護介面，新增甲狀腺疾病判讀報告知識庫及甲狀腺疾病本體論。
2. PAL 推論模組：是本系統邏輯查詢的核心模組，透過 Protégé 2000，甲狀腺疾病影像判讀報告知識庫及甲狀腺疾病本體論可針對使用者的查詢式進行邏輯推論，找出符合語意所需之報告。
3. Query Tab 查詢模組：是系統中圖型化的查詢模組，提供使用者固定的問句並負責執行 Query Tab 查詢介面中的查詢式，透過甲狀腺疾病判讀報告知識庫及甲狀腺疾病本體論，可找出滿足使用者查詢式的報告。
4. 數值型模組：是系統用來處理不同長度計量單位存儲的模組，透過甲狀腺疾病本體論中的單位知識，可找出符合使用者需求的單位數值。數值型模組可以與 Query Tab 查詢模組及 PAL 推論模組整合並提供單位查詢的支援。
5. 使用者查詢介面：分成兩個查詢子介面 PAL 查詢介面、Query Tab 查詢介面。PAL 查詢介面是使用者輸入 PAL 查詢式的介面，提供 PAL 邏輯表示法的編輯器，並將使用者輸入的查詢式送到 PAL 推論模組進行解析。Query Tab 查詢介面是系統中圖型化的查詢系統，它提供使用者固定的問句，使用者填入問句中空白的欄位後，它會將問句傳給 Query Tab 查詢模組以找出想要的報告。
6. 知識庫維護介面：是 Protégé 2000 內定的知識庫維護介面，它可用來新增甲狀腺疾病判讀報告知識庫及甲狀腺疾病本體論於 Protégé 2000 的知識庫之中。透過 Protégé 2000 的動態介面產生器，知識庫維護介面可以依照知識工作者的需求加以調整。

7. 甲狀腺疾病判讀報告知識庫：是系統知識庫的一部份，儲存醫院編碼後的報告。甲狀腺疾病本體論也是整個系統所要查詢的標的物，它的編碼方式是依照甲狀腺疾病本體論中所提供的知識模組來組織及管理報告。
8. 甲狀腺疾病本體論：是系統的知識庫主體，它包括甲狀腺疾病判讀報告本身知識的分析，及甲狀腺疾病判讀報告品質管理運作與報告的相關知識。

二、PAL 推論模組

以本體論為主的知識管理系統中，推理機扮演著非常重要的角色，它提供了啟發式（heuristics）的解答機制，並讓系統了解本體論中的知識以回答使用者的查詢。本研究在 PAL 系統中加入 Axiom Evaluator 機制，使整合後的 PAL 系統能扮演推理機的角色，並回答使用者的查詢。一般專家系統中最常用的推理策略有兩種：向前結合（Forward chaining）及向後結合（Backward chaining）。在 Forward chaining 的情況下，推理機由已知的事實出發，尋找該事實所能適用的法則，直到找出所需結論為止。Backward chaining 與 Forward chaining 的方向恰好相反，是一個以目標為導向的推論方式，它先決定所要追求的結論，然後再尋找資料來驗證所認定之結論的正確性（梁定澎 民 93）。PAL 推論模組的推論機制即是採向後結合的推論技術。

三、數值型模組

影像判讀報告內容中所記載之數值，大多與腫瘤有關，對後段之治療與處置的影響極其深遠（例如針對腫瘤的處置方式為開刀、服藥控制、化療...等），所以對數值精準度的要求，不許有絲毫的偏差。在本研究中，腫瘤大小的描述，所出現的長度計量單位雖然只有 cm（公分）及 mm（釐米）二種，但對使用者而言，仍有搜尋的困擾，因此，同時由不同的長度計量單位進行搜尋的問題仍需克服。針對此一問題，將藉由數值型模組來分別存放報告內容所記載之 cm 或 mm 二種長度計量單位，並由數值型模組支援 PAL 推論模組或 Query Tab 查詢模組。在此架構下，使用者只需透過 PAL 推論模組或 Query Tab 查詢模組便可找尋到與報告內容所記載之相關的數值，包括 cm 及 mm，或者大於（>）、小於（<）某一數值...等相關之報告。數值型模組架構共分成三大模組。分別為單位知識本體論、數值型模組及 Protégé 2000 系統，茲說明如下：

1. 單位本體論：描述報告中所有單位的知識，提供單位推理機的單位知識。
2. 數值型模組：其運作流程係當資料輸入時，模組先取得來源單位及目的單位，然後數值型模組再分別儲存不同的長度單位，以提供進行單位查詢的推論，最後依使用者需求將所需的結果輸出。
3. Protégé 2000：負責單位本體論及甲狀腺疾病本體論的建制，同時配合數值型模組單位資料的存取。

伍、系統成果

為了能讓系統瞭解甲狀腺疾病本體論報告及其相關知識，第一係將甲狀腺疾病判讀報告建置在 Protégé 2000 知識庫之中，建置的過程可分成三階段：Class 建立階段、Slot 建立階段及 Axiom 建立階段；第二為甲狀腺疾病判讀報告中使用的術語、階層架構、及術語間相關的關聯，轉成系統能瞭解的述語邏輯；第三為一般查詢問句 (Formal Competency Question)；第四即為測試結果，最後為研究討論。

一、Class、Slot 及 Axiom 各階段的建立成果

本研究之個案為南部某醫學中心，並進行搜集九十二年十月一日至九十二年十二月三十一日甲狀腺疾病患者判讀報告中，臨床醫師經由問診並針對患者疾病分類資料所下之診斷碼為 #139 MALIGNANT NEOPLASM of THYROID GLAND 的影像判讀報告，再由其中挑選對影像特徵描述較多的檢查項目作為測試標的，其中以電腦斷層檢查為首、超音波檢查報告與一般胸部檢查次之，總數共 274 筆進行分類及測試。判讀報告內容為根據醫師習慣之用語進行分類，再根據分類及測試結果進行後續本研究之判讀報告文件分析，茲將步驟整理如下：

1. 關聯器官分類：甲狀腺疾病較易影響（侵犯）之器官，如氣管……等。
2. 器官特徵：甲狀腺疾病可能導致某些器官功能受到影響，如聲帶發聲困難；甲狀腺腫瘤壓迫（或切除）後某些器官功能受影響或喪失。
3. 腫瘤特徵（呈現的位置）：報告內容記載腫瘤所在的位置可能是接近、環繞、嵌入……等不同的狀況接近某器官；影像所呈現的腫瘤可能是實心、有水樣物質包含其中……等。
4. 語意分類（用法不同但語意相同）：將諸如「air way compression」與「air way divide」……等「用法不同但語意相同」者劃分為同一類別。
5. 語意分類（用法相同語意相同）：將諸如 malignant neoplasm of thyroid gland……等大多數醫師「用法相同語意相同」者歸於同一類。
6. Terminology gap 標準問題與作法：針對就現狀常用專有名詞「用詞相同語意相同」者，延續使用；「用法不同語意相同」者，建置 Metadata 格式化資料庫系統。
7. 設定不同的長度計量單位：分別以 cm 及 mm 設定為長度輸入欄位。

本研究經由所搜集的判讀報告資料，依前述第三章節之影像判讀報告知識庫系統分類方式逐一建立於 Protégé 2000 系統之中，共建立了 64 個 Class、115 個 Slot、276 個 Instance。為了將研究儘量融合實務作業需求，在 Protégé 2000 系統建立的同時，即將目前醫院內部所使用且與本研究有關的編碼對應代號建立於系統中，除縮短研究時間外，使用者可快速瞭解系統所能提供的各項功能，同時在系統中保留適度彈性，若發現類別或次類別需要異動時，可彈性予以增減以符合實務需求，詳如圖 8 甲狀腺疾病本體論 Class 所建立的各項成果。

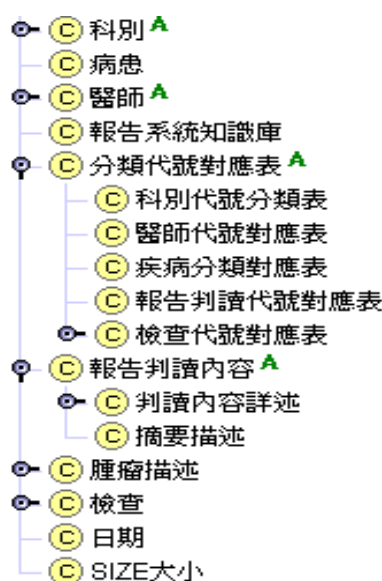


圖 8：甲狀腺疾病本體論 Class

在甲狀腺疾病本體論知識庫中，Slot 代表系統所建立的每一個 Class 所應具備的屬性，整個系統的核心在 Protégé 2000 系統中的報告系統知識庫（此項 Class 居於關鍵位置），系統需經由此項 Class 與其他類別的 Class 產生相互關聯來推論知識之間的關聯性。因此，報告系統知識庫本身的 Slot 需具備與其他 Class 擁有互為反向關聯的聯結能力，此一互為反向關聯如同二個綱要間的 joint（連結）一般，將所有 Class 與 Class 之間的關聯架構起來，此一反向關聯的 joint 在系統中稱之為 inverse（反向）。其次，報告系統知識庫本身所具備的 Slot 包括病歷號碼、相關 tumor、檢查目的及病程摘要、相關 organ、病患基本資料、判讀摘要敘述、判讀之住院醫師資料、判讀之主治醫師資料、開單醫師資料、檢查項目資料、疾病分類資料、報告相關日期、報告內容敘說等，圖 9 為報告系統知識庫 Slot 的建立成果。

由於報告系統知識庫本身的 Slot 必需具備與其他 Class 擁有互為反向關聯的能力，因此，在二個 Class 之間均需互設 inverse 以達到互為反向關聯的目的。舉例而言，在系統中 tumor finding（class 之一）的 slot 包含了 name、location、density、meta、homogeneous/heterogeneous、calcification / cystic change、well/ill define、size 大小、相關部位描述、腫瘤描述...等。為達成 tumor finding 與報告系統知識庫的雙向聯結，在 tumor finding 的 slot 中需設立一 inverse 項目「tumor 相關」作為與報告系統知識庫反向關聯的 joint。而報告系統知識庫則設立一 inverse 項目「相關 tumor」與 tumor finding 建立反向關聯 joint。

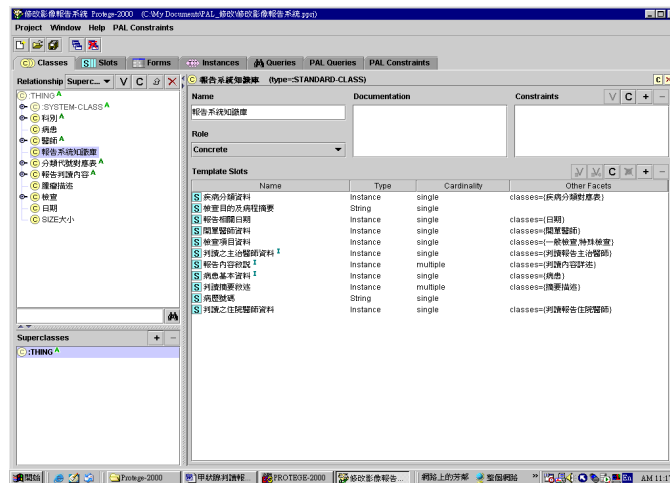


圖 9：報告系統知識庫 Slot 建立成果

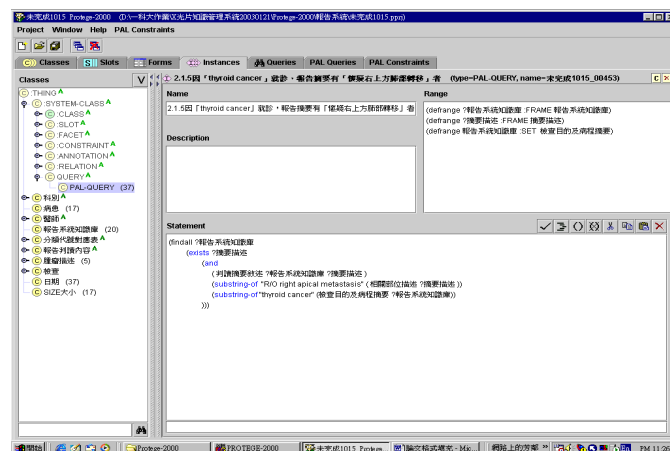


圖 10：「因 thyroid cancer 就診且懷疑右上方肺部轉移」Axiom 建立成果

在人與系統溝通方面，必須透過一些查詢語言讓使用者與機器進行溝通。簡單來說，Ontology 用來描述知識，而查詢語言則是用來推論知識庫內的知識以回答使用者。PAL 是 Protégé 2000 下描述 Axiom 的語言，在 Protégé 2000 的定義下，PAL 是用來表達知識庫中複雜的限制及提供邏輯查詢的語言，而使用的知識模組是採用 Protégé 2000 的架構，所以在執行上能夠與 Protégé 2000 的知識庫完全整合。Axiom 建立階段係將甲狀腺疾病判讀報告中的關聯透過 PAL 邏輯語法推論方式，搜尋使用者所需具有疾病特徵的判讀報告資料，如圖 10 即為搜尋患者因「Thyroid Cancer」就診，報告摘要記錄「懷疑右上方肺部轉移者」Axiom 模式。

二、判讀報告中的述語關聯 (Predicate Term)

以下為甲狀腺疾病判讀報告中所用到的述語關聯：

PT-1：判讀報告_包含_病患基本資料

甲狀腺疾病患者的基本資料屬於判讀報告之部份，二者間互為正反向關聯。

PT-2：判讀報告_包含_檢查目的及病程摘要

病患就診檢查目的及病程摘要屬於判讀報告中部份，二者間為正反向關聯。

PT-3：判讀報告_包含_判讀之主治醫師記錄

只要該醫師隸屬於判讀報告之主治醫師，則判讀報告中必有其判讀之記錄，二者間互為正反向的關聯。

PT-4：判讀報告_包含_報告內容敘說

判讀報告知識庫存檔了醫師判讀記錄，二者間互為正反向的關聯。

PT-5：判讀報告_包含_報告相關日期

所有與判讀報告有關的日期全登錄於判讀報告中，二者間互為正反向關聯。

PT-6：判讀報告_包含_疾病分類資料

疾病分類資料屬於判讀報告的一部份，二者間互為正反向的關聯。

PT-7：Substring – of (a , b)

當字串 a 在字串 b 中出現時，此述語就會成立。

Pred-1. { < 任務關鍵字 > / < 病患基本資料 > }

Pred-2. { < 任務關鍵字 > / < 檢查目的及病程摘要 > }

Pred-3. { < 任務關鍵字 > / < 病灶(Lesion) > }

Pred-4. { < 系統代號類別 > / < 檢查項目代號 > }

Pred-5. { < 病灶(Lesion) > / < 腫瘤 SIZE 描述 > }

Pred-6. { < 判讀報告之主治醫師 > / < 國科會研究計劃(廠商贊助研究計劃) > }

三、一般查詢問句 (Formal Competency Question)

系統除了要瞭解甲狀腺疾病判讀報告及其相關知識外，還須瞭解使用者所會用到的查詢問句 (Competency Question)，Competency Question 中包含有相關的術語及關聯：任務關鍵字、相關的檢查目的及病程摘要、報告內容敘說...等。經由述語的定義，Competency Question 即可用 First Order Logic 來轉換成系統能瞭解的正式化表達格式。以下例舉 Competency Question 的正式化表示法：

CQ1：當使用者給予系統某個任務關鍵字（如：thyroid gland ectomy）時，有那些檢查目的及病程摘要會與這個任務關鍵字有關？

?P 檢查目的及病程摘要與任務關鍵字相關 (?P, "thyroid gland ectomy")

CQ2：當使用者給予系統某個任務關鍵字（如：vocal cord palsy）時，有那些病灶 (Lesion) 會與這個任務關鍵字有關？

?P 病灶 (Lesion) 與任務關鍵字相關 (?P, "vocal cord palsy")

CQ3：當使用者給予系統某個任務關鍵字（如：Tumor > 0.4x 0.5 x 0.6 cm）時，有那些病灶（Lesion）會與這個任務關鍵字有關？

?P 病灶（Lesion）與任務關鍵字相關（?P, "Tumor > 0.4x 0.5 x 0.6 cm"）

CQ4：當使用者給予系統某個判讀報告之主治醫師（如：黃美元）時，有那些國科會研究計劃（廠商贊助研究計劃會）與這個判讀報告之主治醫師有關？

?P 判讀報告之主治醫師與國科會研究計劃或廠商贊助研究計劃

（?P, 「國科會研究計劃或廠商贊助研究計劃」）

Axiom 這個階段是針對 Competency Question 的推論邏輯採用上述述語及其關聯加以定義，以下提供上述 Competency Question 的推論邏輯：

Defn-1：檢查目的及病程摘要與任務關鍵字相關（R, K）

只要檢查目的及病程摘要 R 出現，其任務關鍵字 K（檢查目的及病程摘要代表病況中的「原因」）即成立：

($\forall$ P：報告系統知識庫

($\wedge$ ($\exists$ R：檢查目的及病程摘要 (R, K))))

(findall ?報告系統知識庫

(exists ?摘要描述

(exists ?日期

(and

(判讀摘要敘述?報告系統知識庫?摘要描述)

(substring-of "suspect right apical metastasis" (相關部位描述?摘要描述))

(substring-of "thyroid cancer" (檢查目的及病程摘要 ?報告系統知識庫))

(報告相關日期 ?報告系統知識庫 ?日期)

(or(>(年 ?日期)2001)

(=(年 ?日期)2001))

(or(>(年 ?日期)2004)

(=(年 ?日期)2004))))))

由 <報告系統知識庫 / 報告相關日期 / 檢查目的及病程摘要 > 搜尋與 thyroid cancer 有關聯者。

Defn-2：Tumor Finding 所記載之長度記錄（ $\geq$ 或 $\leq$ ）與任務長度記錄相關（G, K）

只要甲狀腺疾病判讀報告中存在一個 Tumor Finding，其 Tumor Finding 所記載之長度記錄（K=0.4x 0.5 x0.6 cm）是隸屬於報告判讀內容，其任務長度記錄 $G \geq K$ 即成立；

($\forall$ P：報告系統知識庫

($\wedge$ ($\exists$ G：Tumor Finding (G, K))))

(findall ?報告系統知識庫

(exists ?判讀內容詳述

(exists ?腫瘤描述



(and(報告內容敘說 ?報告系統知識庫 ?判讀內容詳述)
 (腫瘤相關描述 ?判讀內容詳述 ?腫瘤描述)
 (or(>(長(SIZE ?腫瘤描述)) 0.4)
 (= (長(SIZE ?腫瘤描述)) 0.4))
 (or(>(寬(SIZE ?腫瘤描述)) 0.5)
 (= (寬(SIZE ?腫瘤描述)) 0.5))
 (or(>(寬(SIZE ?腫瘤描述)) 0.6)
 (= (寬(SIZE ?腫瘤描述)) 0.6)))))

由 <報告系統知識庫 / 判讀內容詳述 / 腫瘤描述 / tumor finding 長度記錄
 $\geq 0.4 \times 0.5 \times 0.6 \text{ cm}$ > 搜尋 $\text{tumor} \geq 0.4 \times 0.5 \times 0.6 \text{ cm}$ 之患者資料。

Defn-3：判讀報告之主治醫師與國科會研究計劃或廠商贊助研究計劃相關(Q,Y)

只要甲狀腺疾病判讀報告內容存在判讀報告之主治醫師(Q)與國科會研究計劃或
 廠商贊助研究計劃(Y)，檢查項目(R)...等隸屬於甲狀腺疾病判讀報告知識庫，其關聯
 就成立。

($\forall$ P：報告系統知識庫
 ($\wedge$ ($\exists$ Q：判讀報告之主治醫師與國科會研究計劃或廠商贊助研究計劃(Q, Y))
 (研究計劃包含耗材及藥品(名稱 R))))
 (findall ?報告系統知識庫
 (exists ?日期
 (exists ?特殊檢查
 (exists ?病患
 (exists ?醫師
 (and
 (報告相關日期 ?報告系統知識庫 ?日期)
 (or(>(年 ?日期)2001)
 (= (年 ?日期)2001))
 (and(檢查項目資料 ?報告系統知識庫 ?特殊檢查)
 (代號 ?特殊檢查 "S0120"))
 (and (病患基本資料 ?報告系統知識庫 ?病患)
 (or (> (年(出生日期 ?病患)) 1939)
 (= (年(出生日期 ?病患)) 1939))
 (and(判讀醫師 ?報告系統知識庫 ?判讀報告之主治醫師)
 (姓名 ?判讀報告之主治醫師 “黃美元”)
 (國科會研究計劃或廠商贊助研究計劃 ?黃美元 “CC002”))
))))))

由 < 報告系統知識庫 / 判讀報告之主治醫師 / 國科會研究計劃或廠商贊助研
 究計劃“CC002”/ 特殊檢查 / S0120> 找出由黃美元醫師所主導的研究計劃，同時該計
 劃與患者所作之檢查項目 S0120 有關聯。

四、測試結果

甲狀腺疾病判讀報告文件的層級模式，將文件大綱概分為：病患資料、檢查目的及病程摘要、檢查項目或疾病分類資料、檢查判讀內容。有鑑於研究的重心在檢查判讀內容特徵的搜尋，在系統架構即依任務需求建立各階段 Class、Slot 及 Instance。本階段將搜尋方向分為四個主要的搜尋類別，明確定義執行搜尋路徑，同時亦藉以延伸 Query 範疇，茲將搜尋方向的主要類別分別說明如下：

1. 以檢查項目為主，可能含括病患性別、年齡、檢查日期區間……等。
2. 病患因某種需求而就診（已具備某種甲狀腺疾病的特徵）。
3. 判讀報告內容具有二種以上特徵（例如患者因 thyroid cancer 而就診）。
4. 不同長度計量單位的搜尋（例如 tumor $\geq 0.3 \times 0.4 \times 0.5 \text{cm}$ ……等）。

Query 語言（採用 PAL 語法）最主要的功能是在既定範圍內執行操作，將個人思想上所希望得到的答案透過 Query 語言來達成，經由 Query 語言的邏輯推理，系統依循已明確定義的路徑，搜尋符合語意之甲狀腺疾病影像特徵（pattern）報告資料並逐一勾勒呈現知識相互間的關聯。因此，當系統內部所架構的判讀報告知識層級愈多且類別愈細時，能搜尋的路徑也變得愈繁複，所能搜尋到的資料內容愈能符合使用者需求。茲將甲狀腺疾病判讀報告中用到的述語關聯與 Formal Competency Question 併入系統測試，範例詳述如下。

（一）測試範例一

採樣區間：93/1/1 ~ 93/6/30

採樣條件：南部某醫學中心之甲狀腺疾病判讀報告，其疾病分類資料為 Malignant Neoplasm of Thyroid Gland 且較具影像特徵者，以電腦斷層檢查為首、一般胸部檢查為次，共挑出 50 筆，進行測試。

特徵：年齡在 65 歲以下，且於 2001 年以後曾作過 Head and Neck CT 及 Chest CT 之患者。

結果：經由 PAL 的推論及搜尋，符合年齡在 65 歲以下，且於 2001 年以後曾作過 Head and Neck CT（頭頸部電腦斷層）及 Chest CT（胸部電腦斷層）之患者（6 筆），如圖 11、圖 12 所示。



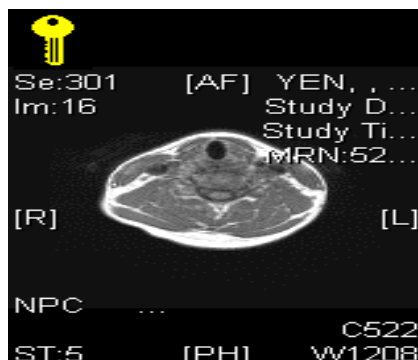
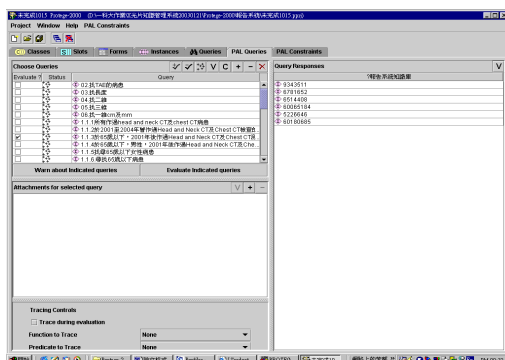


圖 11：65 歲以下病患於 2001 年以後作過 Head and Neck CT 及 Chest CT 者

(二) 測試範例二

採樣區間：93/1/1 ~ 93/6/30

採樣條件：南部某醫學中心之甲狀腺疾病判讀報告，其疾病分類資料為 Malignant Neoplasm of Thyroid Gland 且較具影像影像特徵者，以電腦斷層檢查為首、一般胸部檢查為次，共挑出 50 筆，進行測試。

特徵：因 thyroid cancer 而就診，懷疑右上方肺部有轉移者。

結果：經由 PAL 的推論及搜尋，右上方肺部有轉移者（1 筆）（圖略）。

(三) 測試範例三

採樣區間：93/1/1 ~ 93/6/30

採樣條件：南部某醫學中心之甲狀腺疾病判讀報告，其疾病分類資料為 Malignant Neoplasm of Thyroid Gland 且較具影像影像特徵者，以電腦斷層檢查為首、一般胸部檢查為次，共挑出 50 筆，進行測試。

特徵：判讀報告記載不同的長度單位 tumor size>0.4cm 或>40mm。

結果：經由 PAL 的推論及搜尋，tumor size>0.4cm 或>40mm (5 筆)（圖略）。

五、討論

經由上述研究測試得知，傳統資料庫僅以 keyword 為搜尋要件無法滿足醫療影像研究上的實務需求，物件資料庫以「階層」為分類原則有其困境以及 SQL Sever 無法以語意完整分類。因此，以 Ontology 相關技術解決「診查報告」的判讀問題較佳，茲詳述如下：

1. 傳統資料庫無法滿足情境上的需求：腫瘤可能瀰漫、擴散、接近、嵌入在某些（個）器官，倘以 keyword 搜尋僅能針對某一器官或某一狀況（以交集方式搜尋；例如以【接近】……等單一字樣查詢）。

2. 物件資料庫階層式設計的困境：階層式的設計係以器官為主，將某器官受影響後可能出現的狀況逐一設計於系統中，惟此法可能設限病灶僅能發生（影響）在某器官，當受影響的器官範圍擴大時，則無法適用。例如癌細胞轉移時，可能同時有二個以上的器官受影響。
3. SQL Sever 無法以語意完整分類：關聯式資料庫必需是可明確劃分階層歸屬，亦即判讀報告用語需備有一致性始可交互關聯，在醫學用語尚無法趨於一致以前，較難呈現完整的搜尋結果。

陸、結論及研究方向

一、結論

對醫療服務業而言，儲存判讀報告是累積知識的方法之一，但隨著時間的流逝，報告的數量也會越來越多，在此種環境下，報告管理是醫療服務業的一條必經之路。以往報告管理系統常因報告量過大且報告內容複雜，導致系統無法找出使用者真正想要的報告，本研究針對醫療服務業所面臨的問題，利用本體論的技術，在醫療服務業中建立起一個具推論式的報告知識管理系統。在系統中，使用者可仰賴系統的推論機制以找出真正想要的報告，不須以人工重覆過濾報告的內容。動態的知識仍需要適時更新與維護，才能延伸知識的領域。甲狀腺疾病影像判讀報告本體論知識系統嘗試在特定的疾病上動態呈現判讀報告醫師內隱的知識，透過系統的引領與探鑽，研究人員可再深入探究疾病特徵彼此間的關聯，更可藉此發展研究模式與假說。

本研究的貢獻茲摘要如下：

1. 利用本體論的技術，建置一個能瞭解語意的甲狀腺疾病影像判讀報告知識庫結構，藉由知識庫的結構得以劃分影像判讀報告中不同的知識層級，逐步架構起整份報告所隱含的疾病因果關係，使得專家的隱性知識變成顯性知識。
2. 相較於傳統資料庫，此一知識系統最特別之處在於能「語意分析專家所突顯強調的病灶特徵」，而非 keyword，使得專家知識的呈現更趨完整。
3. 藉由語意觀念深入甲狀腺疾病影像判讀報告的知識分析與系統實作，來達到語意搜尋的目的，以提高研究人員對影像特徵的搜尋比率，及掌握及腫瘤大小與病灶轉移之間的關聯性，縮短研究資料搜尋時間。

二、未來研究方向

針對本研究未來可繼續執行的研究方向及重點，歸納出二個主要的方向：

（一）分析描述某物件的功能、物件轉移現象及物件之間的規則

在未來可繼續執行的研究方向及重點：分析描述某物件（器官）的功能、物件轉

移現象及物件之間的規則，即物件（器官）本身具有某些功能，當某些功能喪失或產生障礙時會形成某些現象，彙集這些現象即可成為規則。在報告分析的過程中，若能提供某些治療或處置方式會對病況的改善有所助益，且可明確說出其可行性百分比供病患或家屬參考，為未來一大努力方向。

（二）自動邏輯說明的能力

在進行本體論工程時，本研究是透過知識工作者的協助將推論邏輯分析出來。倘若在醫療專業領域上，經由疾病專門字彙的使用，可連結至相關字彙的 Ontologies，會更便於萃取相關知識。因此，如何讓疾病知識地圖的表達，能透過圖解範例的呈現更趨完整，亦是未來一大努力方向。

參考文獻

1. 王德玲譯、漢彌頓·畢禮等著、民 92，延續管理，台北：天下雜誌。
2. 何美玉，1996『”ISO 認證：化『礙手』為『順手』』，管理雜誌，第 270 期：44～47 頁。
3. 吳昭新、李友惠等著，2002，醫療資訊管理學，台北：偉華書局。
4. 林郁馨，民 90，口腔超音波網路教學系統之建立與評估，私立台北醫學大學口腔復健醫學研究所碩士論文。
5. 林素蓉譯、Davi-Ellen Chabner 著，2000，醫護術語，台北：華騰文化。
6. 胡訓誠，民 92，應用本體論設計 ISO 文件管理資訊系統，國立高雄第一科技大學資訊管理研究所碩士論文。
7. 范碧玉，民 89，病歷管理理論與實務，台北：合記書局。
8. 范振添，民 91，醫療影像傳輸及資料探索之系統開發，私立中原大學醫學工程研究所碩士論文。
9. 梁定澎，民 93，決策支援系統與企業智慧，台北：智勝文化。
10. 楊宏智、林宏達，1995『甲狀腺機能亢進之治療』，臨床醫學，第三十六卷·第三期。
11. 經濟部標準檢驗局，1999，NII 資訊及通訊國家標準應用推廣手冊-遠距醫療相關標準。
12. Bates, S. E., Meadows, B., Goldspiel, B. R. et al., “A Pilot Study of Amiodarone with Infusional Doxorubicin or Vinblastine in Refractory Breast Cancer”, *Cancer Chemother Pharmacol* (35) 1995, pp:457-463
13. Bechhofer, S. & Goble, C., “Thesaurus Construction through Knowledge Representation”, *Data and Knowledge Engineering* (37:1) 2001, pp25-45
14. Gennari, John H., Musen, Mark A., Ferguson, Ray W., Grosso, William E., Crubézy, Monica, Eriksson, Henrik, et al., “The Evolution of Protégé: and Environment for

- Knowledge-Based Systems Development”, *International Journal of Human-computer studies* (58:1) 2003, pp:89-123
15. Gupta, A., Ludäscher, B., Grethe, J. S., & Martone, M. E., ”Towards a Formalization of Disease-Specific Ontologies for Neuroinformatics”, *Neural Networks* (16:9) 2003, pp:1277-1292
 16. Jacquelinet, C., Burgun, A., Delamarre, D., Strang, N., Djabbour, S., Boutin, B., Le Beux, P. et al.. “Developing the Ontological Foundations of a Terminological System for End-Stage Diseases, Organ failure, Dialysis and Transplantation”, *International Journal of Medical Informatics* (70:2-3) 2003, pp:317-328
 17. Kim, Henry M., “XMLhoo!: A Prototype Application for Intelligent Query of XML Documents using Domain-Specific Ontologies”, Proceedings of the 35th Hawaii International Conference on System Sciences(HICSS-35’02), USA 2002.
 18. Ridley, E. L., “PACS News: ACR/NEMA to standardize DICOM file Management”, *Diagnostic Imaging*, 1995, pp:5-9
 19. Staab, S., Angele, J., Decker, S., Erdmann, M., Hotho, A., Maedche, A. et al. “Semantic Community Web Portals ”, *Computer Networks* (33:1-6) 2000, pp:473-491
 20. Tamma, V., Phelps, S., Dickinson, I., Wooldridge, M., “Ontologies for Supporting Negotiation in E-commerce ”, *Engineering Applications of Artificial Intelligence* (18:2) 2005, pp:223-236
 21. Uschold, M., “The Use of the Typed Lambda Calculus for Guiding Naive Users in the Representation and Acquisition of Part-Whole Knowledge”, *Data & Knowledge Engineering* (20:3) 1996, pp:385-404
 22. Weng, Sung-Shun, Tsai, Hsine-Jena, Liu, Shang-Chia, Hsu, & Cheng-Hsina, “Ontology Construction for Information Classification”, *Expert Systems With Applications* (31:1) 2006, pp:1-12
 23. White, Richard H., Brickner, Leslea A., & Scannell, Kate A., “ICD-9-CM Codes Poorly Identified Venous Thromboembolism during Pregnancy”, *Journal of Clinical Epidemiology* (57:9) 2004, pp:985-988

