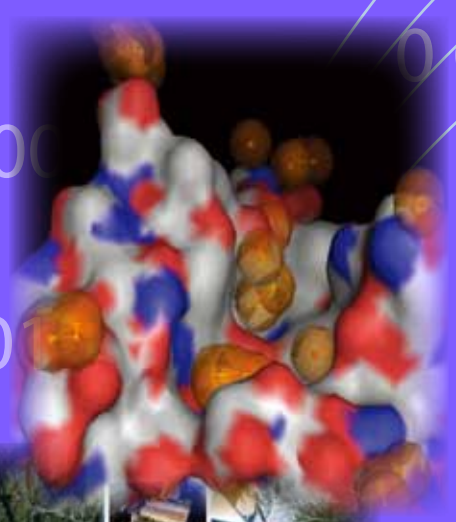
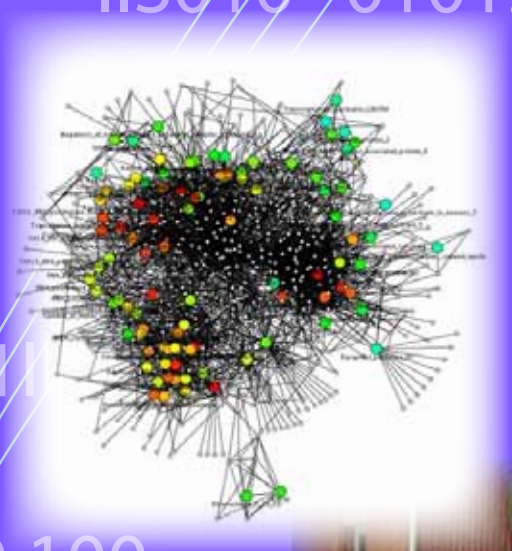
2010 創刊號

IIS Update

資訊科學簡訊

Newsletter of the Institute of Information Science
Academia Sinica, Taiwan





焦點報導

2 所長的話
游本中
3 資訊所名人榜
4 實驗室報導
生物資訊實驗室
以生物資訊探索
生命奧秘
5 卓越講座系列
6 活動快訊
8 計劃特寫
玉山網守護愛登山
的您
9 計劃特寫
人工智慧突破：
果蠅「腦地圖」
10 人物焦點
從程式人跨界學術人
陳昇瑋
11 創意發想
再看二元搜尋法

Institute of Information Science
Academia Sinica

Institute of Information Science

資訊科學研究所於 1977 年開始設立籌備處，歷經五年籌備，於 1982 年 9 月正式成立研究所，是中央研究院數理組十個單位之一。本所的研究重點分為八大方向，包含生物資訊實驗室、電腦系統實驗室、資料處理與探勘實驗室、多媒體技術實驗室、語言與知識處理實驗室、網路系統與服務實驗室、程式語言與形式方法實驗室、計算理論與演算法處理。

本所並於 2002 年配合中研院國際研究生院 (TIGP)，開始招收生物資訊學程博士班，致力延攬及培訓此一新領域之國際人材，七年來已招募 31 名學生，對生物資訊的研究有很大的助益。

所內資深研究人員皆指導或共同指導來自全國各地頂尖大學之碩、博士生，並與這些大學保持緊密的研究關係及合作計劃，與國內外資訊業界亦有極為良好之互動關係。

所長：游本中

副所長：高明達
　　陳克健

研究群主持人：
宋定懿教授
　生物資訊實驗室
游本中教授
　電腦系統實驗室
陳孟彰教授
　資料處理與探勘實驗室
廖弘源教授
　多媒體技術實驗室
許聞廉教授
　語言與知識處理實驗室
何建明教授
　網路系統與服務實驗室
莊庭瑞教授
　程式語言與形式方法實驗室
高明達教授
　計算理論與演算法處理

所長的話

中央研究院是台灣最高的學術研究機構。本所成立於 1982 年，致力於資訊科學及其相關技術之基礎及應用之研究。目前約有 40 位研究員，22 位博士後研究學者，以及 300 多位專任之資訊技術人員與博士生。從事於各種不同領域的研究主題，範圍廣泛，從生物資訊、多媒體技術、網路系統與服務，到中文自然語言與知識處理、計算理論與演算法。

身為中央之研究單位，資訊所經由各實驗團隊跨學科的研究計劃，著力於各項中長期的基礎研究及其應用之相關議題，試圖與台灣各大學研究單位或工業研究單位區隔，彼此間的不同。

本資訊科學簡訊創刊號以及之後每半年發行一期的簡訊刊物，我們誠摯的邀請您分享本所精彩的活動以及各項卓越的研究。期許發掘默默耕耘投入的人，以及他們的最新訊息，並從各種不同的研究視野，探索一些有趣的點子及創意。

本簡訊內容為了達到豐富多樣的目標。首先，我們提供了資訊科學研究社群與一般大眾間，一個交流的平台，尤其是海外讀者，希望讓他們了解到我們在研究上的努力及成果。資訊所名人榜表彰了他們專精的研究成果，在海內外獲致各項榮譽及獎項；此外還包括我們舉辦的各項活動訊息。同時也提供了落腳在各方的資訊所校友們，一個連繫的機會，藉此得知他們的近況與消息。因此，我們鼓勵他們投稿分享他們最新的消息與成就。

我們也希望透過簡訊的鼓勵，召募到更多優秀的人才加入我們，成為我們團隊的一份子。的確，我們希望本簡訊成為各方意見、訊息及資訊交流，一個有效的平台。因此我們竭誠的歡迎您的投入，提出您的意見及貢獻。

名人榜

- 特聘講座張韻詩教授獲頒第七屆 Linux 黃金企鵝特殊貢獻獎
- 所長游本中教授榮獲 IICM「榮譽獎」
- 陳昇瑋助研究員榮獲 ACM 台北 / 台灣分會「李國鼎青年研究獎」
- 陳銘憲特聘研究員榮獲 98 年度行政院國家科學委員會傑出研究獎
- 許聞廉特聘研究員榮獲潘文淵文教基金會 2010 研究傑出獎
- 楊得年助研究員榮獲潘文淵文教基金會 2010 考察研究獎

副研究員楊柏因先生、助研究員楊得年先生榮獲本院 99 年度前瞻計畫

高效能密碼計算

此計畫旨在確認各種密碼系統 -- 和對它們的攻擊 -- 可以實做得多快，從而幫大家提高現實生活中的安全性和隱私。現代密碼學的宗旨是：有法固有破，只要足夠的金錢和時間，即使決策者總是聲稱安全性比效能重要，仍不免需在兩者間做取捨。實做上，加密還經常比破密要求更高的效率。例如，今天網路商業能發展都因為大家能信任 RSA -- 和其他幾個 -- 密碼系統。RSA 的安全性則取決於 RSA 本身與攻擊它的演算法 -- 例如一般數體篩法 (GNFS) -- 可以實做得多快。

相反地，更有效的加密方法，可以讓我們在相同的價格得到更好得安全性，或是在相同的安全性下使用較便宜的組件。RSA 的計算量太大，不適合低資源的場合例如全在計算。北約組織因此也將揚棄 RSA 標準而易以橢圓曲線密碼。放眼將來，研究的主力也會放在橢圓曲線，以及不會被未來的量子電腦擊倒的密碼系統 -- 所謂的後量子密碼演算法 。為了實行本研究計畫，必須瞭解許多不同的算法，並在各種計算機結構下作最佳化。此團隊對此頗有經驗 -- 不管是在一般電腦，微控制器或是顯示卡上 -- 也因此自 2004 以來在國際密碼學會 (IACR) 的各研討會上有不

(文轉第 9 頁)

網際網路資料儲存應用之選擇性網路編碼技術

本計畫擬提出適用於網際網路資料儲存應用之選擇性網路編碼技術。目前網際網路資料儲存伺服器之作法，是將每一筆資料單獨進行儲存；而本計畫所提出之選擇性網路編碼技術，乃藉由資料混合概念，使伺服器可以儲存更多資料。然而，傳統的網路編碼技術，在設計上主要針對群播與廣播之環境。此外，目前網際網路上的資料流仍以單點傳輸為主，且傳統的網路編碼技術並不相容於目前之路由器以及國際標準。因此，傳統的網路編碼技術，目前仍無法被廣泛使用。因此，本計畫擬提出利用使用者擷取資料之規律，將網際網路儲存內容進行編碼。此作法在實務上較為可行，可完全相容於現有的網際網路相關設備與國際標準。在本計畫中，首先將探討在單一儲存伺服器中，資料之選擇以及編碼方式，以提高資料快取符合率。在傳統儲存應用中，每筆資料只能服務查詢該資料的客戶。相較之下，將藉由多重資料編碼技術，讓每一筆編碼資料可以服務更多使用者。此外，對於線上運算與資料更新之情境，本計畫提出兩個新穎的快取管理機制：彈性存取與適性調整，以增加空間分配之靈活性，並減少資料快取失誤率。此外，更進一步探討其應用於分散式儲存伺服器之機制設計。

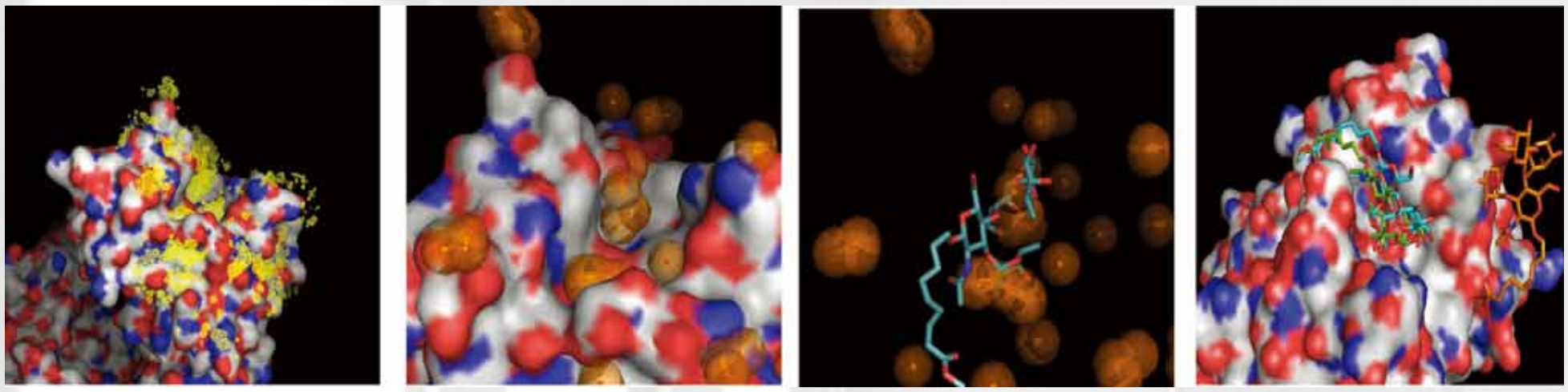
研究員廖弘源先生榮獲本院 99 年度深耕計畫

TELDAP 多媒體影像處理技術

中研院資訊所多媒體技術實驗室近幾年來將研究重心放在兩個主要的領域：（一）多媒體訊號處理；及（二）多媒體相關應用。在多媒體訊號處理方面，其目標是要對此領域有基礎的貢獻。至於在多媒體應用領域，目標是整合過去該團隊成員或別人發展出來的系統，成為對人類真正有用的整合系統。在過去十多年中，多媒體技術實驗室在影像 / 視訊處理、三維電腦繪圖、機器學習及多媒體安全及保護等領域多所建樹，發展許多舉世聞名的技術，例如：雞尾酒浮水印技術、人臉辨識技術及行動中物件追蹤技術等。另外，有一些實作系統已成功技轉給廠商並獲得經濟部及台北國際電腦展（2009 Computex Taipei）的獎勵。目前，多媒體實驗室參與兩個大計劃。一個是由國科會支持的數位典藏及數位學習國家型計畫 (TELDAP)，另一個是由經濟部支持的學界科專計畫 (VBIE)。身為 TELDAP 計畫 R ＆ D 的主持人，廖博士帶領團隊成員為計畫所收集的多媒體資料（包括文字資料、老照片資料，及舊的新聞影片、老電影等）研發一些核心技術，因為這些老舊資料都是數十年前拍攝，加以數位化之後，並發展相關資料壓縮、註解，及快速擷取的技術。因為 TELDAP R ＆ D 分項的經費大致分成三大部份，一部份做標準化及 Metadata 的技術研發，一部份做多媒體系統之技

(文轉第 8 頁)

以生物資訊探索生命奧秘

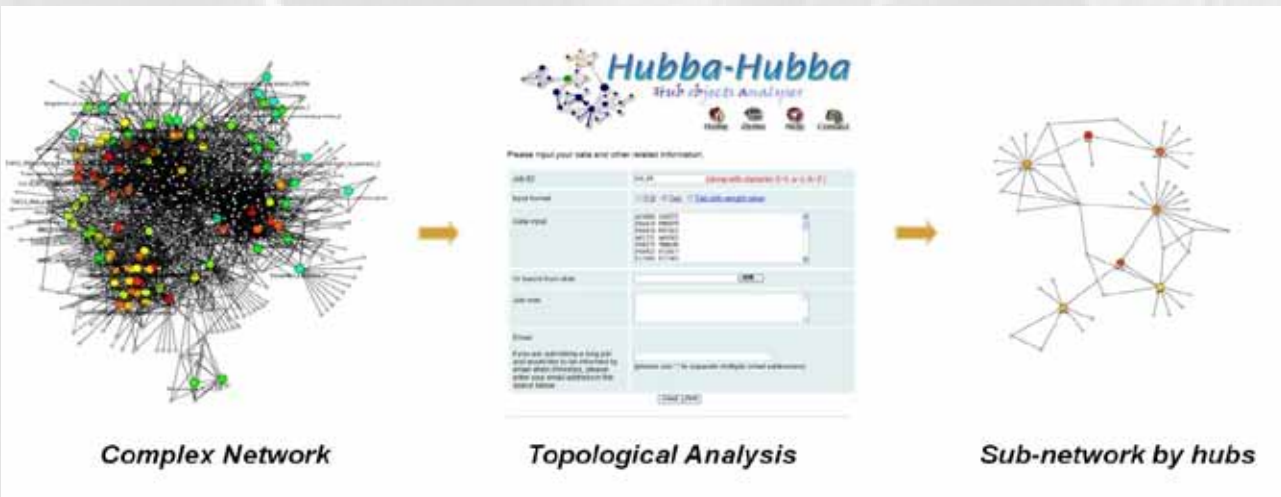


蛋白質與化合物連結預測

生物資訊實驗室的目標在於發展目前在生技研究上迫切需要新的生物資訊系統，尤其希望藉發展的系統找出新的生物發現。

此實驗室目前研究方向大致分為：

（一）系統生物學：包括：1、調控機制與網路：研究酵母菌以及更高物種（如：人類）上基因之調控機制，包括尋找微 RNA 在基因調控扮演的角色、辨識轉錄因子的結合位置、及具退化性質的結合位置對基因表現的影響以及鄰近基因的調控機制之異同。2、網路生物學：由目前大規模實驗所產生的大量生物交互網路，研究乳突病毒與 C 肝病毒在宿主細胞上的交互作用，從中找出複雜網路中的重要次網路與代謝調控模組，再以此為起點，尋找與人類疾病機轉與病原感染機制相關的重要蛋白質，以作為新的藥物研發標的與治療方法的突破關鍵。3、環境基因體學：利用新一代高速定序技術，可將存在於生物體內外與環境中的微生物群體，進行全面性的組成種系鑑定與生化功能分析，並針對不同環境下微生物相發展比對分析平台，以利快速找出特定環境下的微生物或病原體序列及生質能源相關的基因組成。4、轉錄體分析：持續發展非模式物種高通量轉錄體定序的生物資訊服務平台，支援特定本土非模式物種的轉錄體相關研究。這些非模式物種的基因體尚未組合完成，相關演算法的設計也將是一項挑戰。

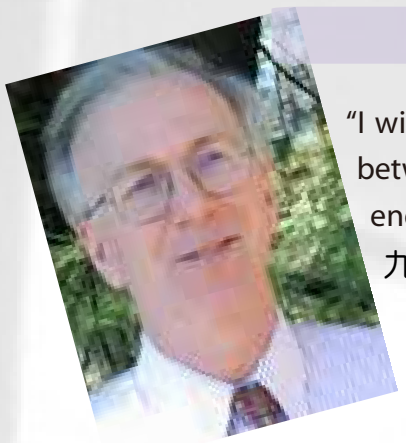


（二）蛋白體學：1. 質譜儀分析的蛋白體學：針對質譜儀實驗產生的大量數據，探討蛋白體學上兩個最重要且根本的問題：蛋白質鑑定與蛋白體定量，以找出在疾病細胞與正常細胞上表現不同的蛋白質為可能的生物標記。2、蛋白質預測：利用機器學習，進行不同蛋白質的預測問題，包括：蛋白結構預測、蛋白質細胞定位預測。有關前者，同時將針對膜蛋白預測，因目前約有 50％的醫學上的生物標記是膜蛋白類。後者與蛋白質功能相關，並且已發展出一套多重細胞位置預測的軟體工具，後續將針對不同物種及胞器之定位預測進行研究。

除此之外，該實驗室並參與大型計畫，如：從纖維素到生質酒精關鍵技術之開發（國科會計畫，計畫總主持人：翁啟惠院長）、C4 水稻計畫（中研院計畫，計畫主持人：李文雄院士）、基因體醫學及生技研發之生物資訊核心設施（基因體醫學國家型科技計畫，計畫主持人：張憶壽教授，參考網站：http://www.tbi.org.tw)，以及生物資源整合與建置與計畫（國科會計畫，計畫主持人：張清風教授（海洋大學副校長）和國科會生物處郭明良處長）。

該實驗室除了進行研究工作外，並於民國 92 年開始國際研究生院生物資訊博士班學程，目前除 2 名畢業外，有 34 名學生在學，包括本國籍，以及歐亞不同國家，如：印度、馬來西亞、越南、菲律賓、德國、美國、加拿大、斯洛伐克。

卓越講座系列



"I will present, AlarmNet, a novel testbed system for health care that uses a two-way flow of data and analysis between front-end body networks, intermediate environmental sensing and communication networks, and back-end context aware protocols that are tailored to residents' individual living patterns." **Dr. John A. Stankovic.** 九十九年一月四日 . **Challenges in Mining Cyber-Physical Systems**

"Cyber-Physical Systems are integrations of computation, networking, and physical processes. Both static and dynamic data, in various types and formats, are generated and collected from the interconnection of cyber elements and physical elements." **Dr. Lionel M. Ni.**, 九十九年三月八日 . **Challenges in Mining Cyber-Physical Systems**

"We will talk about how FTR pursues disruptive technologies, and also highlight several ways Intel Labs engages university and national lab researchers." **Dr. Andrew A. Chien,** 九十九年四月六日 . **Bold, Edgy Technologies for the Future of Computing**

" In this talk, we introduce the broadband network environment in Taiwan, and describe a Next Generation Network (NGN)/IP Multimedia Subsystem (IMS) deployed by Chunghwa Telecom." **Dr. Yi-Bing (Jason) Lin,** 九十九年五月四日 . **Deploying Applications in Next Generation Networks**

2010 年卓越講座系列

七月	Dr. Jun'ichi Tsujii	Tokyo University	Natural Language
八月	Dr. Chin-Hui Lee	Georgia Tech	Speech and Speaker Recognization
九月	Dr. Frans Kaashoek	MIT	Operating and Distributed Systems
十月	Dr. Jean Ponce	Ecole Normale Superieure	Computer Vision
十一月	Dr. Tom Mitchell	Carnegie Mellon University	Machine Learning
十二月	Dr. Bart Preneel	Katholieke Universiteit Leuven	Information Security

~ 最新演講訊息，請上本所網站 ~

參訪團

法國最高研究單位 INRIA 國合處，處長 Dominique Sotteau 及 Emmanuelle Platzgummer 於 12 月 2 日蒞臨本所，希望藉由對本所的深入了解，使台法雙方在資訊領域，達到進一步的合作關係。

北一女中資訊研習社於 4 月 24 日至本所參觀訪問。

伊利諾伊大學香檳分校 ECE 系主任 Andreas Cangellaris 率 Milton Feng 教授，及對外事務主任 Beth Katsinas 於 11 月 6 日至本所參訪及學術交流。

中國華中科大參訪團於 12 月 8 日率團來訪，對本所留下深刻印象。

印尼學校共約二十九位師生於 3 月 23 日來訪，由羅光倫博士介紹本所。

自強活動開跑！

暑假最強檔的自強活動，
提早開跑囉！
5 月 27 日至 5 月 29 日
風光明媚的國境之南
墾丁二日遊。

爬上英國領事館的長梯。

墾丁小灣沙灘。

高雄旗津天后宮。

國際研究生迎新晚會。

國際研究生學程新生訓練。

資訊所尾牙，右圖：舞蹈表演
下圖：摸彩活動

2010 第四屆「邏輯、語言與計算」暑期研習營登場

2010 年「邏輯、語言與計算」暑期研習營於 6 月 28 日至 7 月 9 日正式開辦，今年已邁入第四屆，將以高等程式語言為主題，希望能達到培養學員獨立進行基礎計算科學研究之能力。

近年來，從事基礎計算科學研究之台灣學者們已在各校各自成立研究團隊並相互合作。故特別開設整套課程，邀請相同興趣的學者一同開課，訓練下一代的研究人才。

日期：6 月 28 日至 7 月 9 日
時間：星期一至星期五 9:00am-5:00pm
地點：國立台灣大學進修推廣部

詳細內容請參考：http://flolac.iis.sinica.edu.tw/flolac10/

IEEE International Conference on Service-Oriented Computing and Applications SOCA 09 國際會議於 2009 年 12 月 14-15 日舉行
左至右：林桂傑教授 (大會主席), Raj Rajkumar(主講者), Tharam S. Dillon (主講者), 陳銘憲教授 (CITI 主任), 以及金仲達教授 (議程主席)。

第十六屆高性能計算及嵌入式編譯技術研討會於 2010 年 5 月 27 日至 28 日舉行為期兩天的活動。

玉山網守護愛登山的您！

玉山網是一個針對登山客安全而發展的登山資訊系統

玉山網計畫起緣於 2008 年，由一群結合登山與無線感測網路專長的研究團隊，在玉山國家公園管理處的補助與委託下所共同創立。此計畫的目的是在玉山國家公園內，發展一個高可靠度的登山客追蹤與搜救系統。玉山網目前已正式成為國科會所補助的三年期無線感測網路建置計畫，並將成為玉山國家公司未來的正式服務項目。

玉山網建置於長達 10.9 公里、高度落差達 1302 公尺的玉山主峰登山步道上，並在重要的路段上架設基地台。當登山客由登山口出發時，將被要求配戴一個稱作『黑盒子』的小型裝置，此裝置的重量約 160 公克，並配有衛星定位器、無線收發器與 10KB 的記憶體。當兩個登山客在山徑上相遇時，兩個黑盒子會自動互相記錄彼此的裝置編號，以及相遇的時間和地點資訊，並將所收到的資訊存在記憶體中。當登山客經過基地台時，黑盒子會自動上傳記憶體中所有接收到的資料，而基地台則會將這些資料傳送到遠端的資料伺服器上。

玉山網計畫主要由三個研發團隊組成。其中，在系統佈建方面，由台灣大學黃寶儀教授帶領，旨在 (1) 擴大佈建規模，預計在 2010 年達到 100 個黑盒子，並在 2011 年增加到 200 個黑盒子的規模；(2) 於 2010 年在登山路徑上建置至少三個基地台 (例如：登山口、西峰涼亭、排雲山莊)，並於 2011 年再增建兩個基地台在圓峰山屋與孟祿涼亭；(3) 提供方便使用的網頁介面，供登山客與一般大眾瀏覽、查詢登山客的登山軌跡。在技術研發方面，由中央研究院陳伶志研究員與成功大學藍崑展教授共同領導，任務為 (1) 整合衛星定位器、三軸加速度器與其他相關的感測器於黑盒子中；(2) 設計適合登山環境的感測網路省電演算法，以延長黑盒子的使用時間；(3) 增進系統的正確性、穩定性、可靠性與效能。此外，在創新服務研究方面，由台灣大學朱浩華教授指導，重點在 (1) 提供玉山國家公園詳實的資訊，用來發展先進的旅遊服務；(2) 與玉山國家公園合作開發全新的電子旅遊服務。透過玉山網系統的研發，預期將能 (1) 提供一個真實、可用的無線感測器網路；(2) 獲得設計、開發、佈建與維護大型無線感測器網路的第一手經驗；(3) 給予社會正面的影響，並增加無線感測器網路研究在全球的能見度；(4) 激發學術界與社會對無線感測器網路的興趣。

玉山網的成果將可被延伸到許多領域，例如野生動物追蹤、科學監測、土石流監測與救災網路等。欲了解更多玉山網的資訊，請參觀玉山網計畫網站 http://nrl.iis.sinica.edu.tw/YushanNet/ 或是透過電子郵件聯絡 YushanNet@gmail.com。

玉山地標

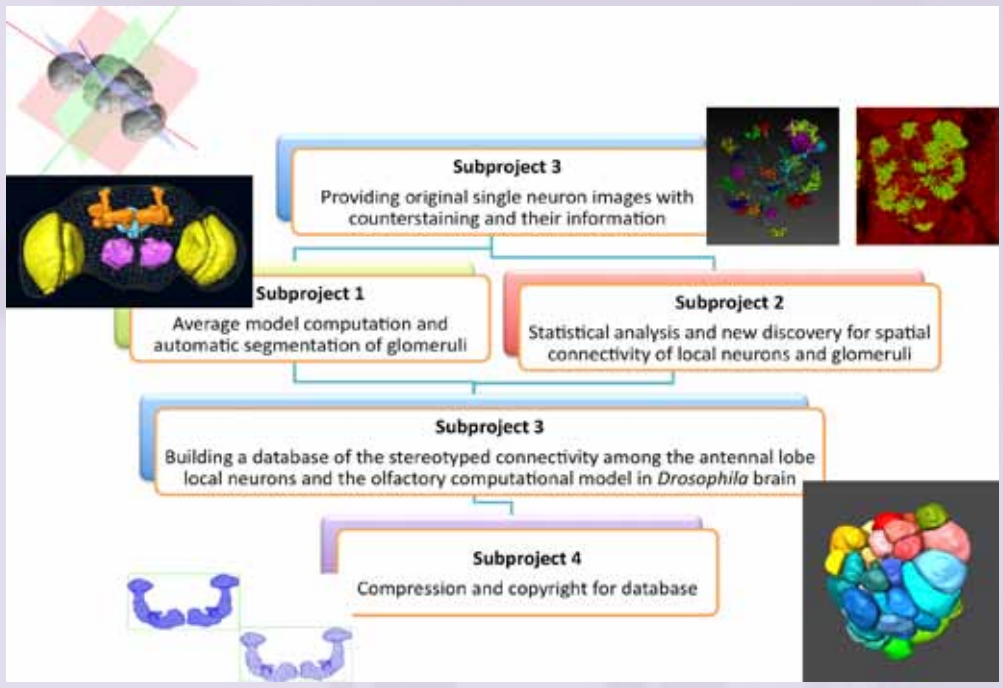
玉山主峰步道

玉山網計劃成員

(轉自第二頁)

術支援，最後一部份才是核心技術研發。因此，在人力配置上有些許不足。尤其是這些核心技術有許多項目都是頭一次做，根本沒有把握做的出來。在本所游本中所長的支持及提名之下，由廖教授獲得深耕計畫之資助，有較多的人力可以投入較前瞻卻難以掌握做出來的 TELDAP 相關核心技術研發工作。

人工智慧突破：果蠅「腦地圖」



清大合作計劃／總計畫主持人：黃文良

具體目標

本整合型研究計畫的目標可概述如下兩點：（一）研究果蠅腦嗅覺葉(antenna lobe)內的計算網路；（二）了解嗅覺葉內部的嗅覺運算機制。因此，為了完成這兩項目標，分為四個子計畫進行研究；在其中，子計畫一與子計畫二偏重於利用影像處理和統計方法，來對相關資料進行處理分析。這兩個子計畫的目的在於分析由子計畫三所提供的高解析度生物醫學影像，並從中萃取出足夠充分的資訊，以便於讓生物學家推想出有意義且合理的嗅覺計算模型——這也是子計畫三的主要目的。最後，這些相關的模型資料可藉由網路呈現、散佈傳輸給所有對此領域有興趣的研究人員，當然，所衍生出的議題就是該如何對這些資料進行適當的認證保護，這也是子計畫四的主要目標。這四個子計畫間的關係，則如下圖所示。

這個跨領域、學科的研究方法將大幅地提高其領先其他國際團隊的機會，也必然將對「腦如何將感受(刺激)化為認知」的運算方式的研究產生重大的影響。

四個子計畫的個別研究主題與計畫主持人等資訊，則如下所示：

1. 子計畫一：果蠅腦嗅覺網路的三維影像處理技術；主持人：清華大學電機系陳永昌教授。
2. 子計畫二：果蠅腦嗅覺網路機率圖的建構與分類；主持人：交通大學統計所盧鴻興教授。
3. 子計畫三：繪製在果蠅腦中處理嗅覺資訊的神經元圖譜；主持人：清華大學生科系：江安世教授。
4. 子計畫四：對果蠅腦嗅覺網路資料庫的視訊處理，壓縮及版權保護；主持人：中研院資科所黃文良教授。

創新與競爭優勢

本團隊的競爭優勢在於擁有全世界最高解析度的果蠅腦影像，子計畫三的主持人以及其團隊已開發出一套新型生物成像(bioimaging)技術，以利於用高解析度的方式取像並重建生物體的三維結構，這項技術的成就對全世界的各研究團隊來說，都是難以望其項背的(Liu and Chiang, 2003)。獨特的技術突破，在對腦部迴路進行三維取像時，擁有精準至可針對單一神經元拍照的極高解析度。對其他在同一領域近進行研究的神經科學家而言，此技術甚至已經成為難以跨越的門檻。本成像技術的初步應用之一，已經讓江教授的團隊發現嗅覺運算的神經立體連結關係、和果蠅腦內涉及形成長期記憶的新結構(Lin et al., 2007; Wu et al., 2007)。

預期成果

本計劃的目標是建立一個果蠅腦嗅覺葉內的嗅覺計算計算模型，換句話說，也是一個能展現神經鏈結關係的「腦地圖」。這地圖包括在嗅覺葉內所有參予嗅覺運算的單元，因此，研究人員可藉此推斷出與果蠅行為相符的嗅覺運算方式；這裡所謂的「果蠅行為」，則泛指了可能的神經迴路活動，或單一果蠅個體所表現的行為模式。就目前所知，該研究若能成功達成目標，這份腦地圖將是史上第一個能描述腦部功能間鏈結的範例，也能滿足人類對人工智能的研究上的一個夢想。

為了實現上述目標，子計畫一將利用影像處理的技術，對嗅覺葉內的嗅小球(glomeruli)影像進行影像切割，並建立標準平均模型。與此同時，子計畫二將以統計的方式分析從單一神經元影像中所萃取出的相關參數，並以機率地圖(probability map)的方式來表達可能的鏈結情形。隨後，子計畫一、二將共同建立果蠅腦嗅覺葉內嗅小球與相關神經元的聯接地圖；子計畫三則負責分析來自腎小球和區域神經元(local nuron)連接關係的資訊，並提出一個果蠅腦內嗅覺計算網路的初步模型。最後，這些罕見、珍貴的資料將會以資料庫的方式來對外呈現，資料的傳輸、保護、壓縮等相關處理，則有賴於子計畫四、三之間的溝通與合作。

(轉自第二頁)

錯的成績。事實上，在密碼學硬體與嵌入式系統研討會(CHES)和快速軟件加密研討會(FSE)是唯一的台灣代表，也是除了資訊所研究員呂及人博士以外，唯一登上每年密碼學三大會的台灣團隊。

因為在顯示卡上實做的代數攻擊而獲得 CHES 2010 錄取，這個成果必需感謝前瞻計劃，該團隊也希望在可見的將來繼續因為實做的高效能而得到尊重。

從程式人跨界學術人的深度歷險

陳昇瑋博士目前為中央研究院資訊科學研究所以及資訊科技創新研究中心助研究員，同時是多媒體網路與系統實驗室主持人。他的研究焦點著重在網路效能、使用者滿意度、網路安全、線上遊戲等領域，在線上遊戲及網路使用者滿意度的量測及管理方面持續有代表性的研究創見，更經常在知名的國際會議，包括 ACM SIGCOMM, IEEE INFOCOM, ACM Multimedia, ACM CHI, ACM NetGames, ACM NOSSDAV 等會議發表論文，提出獨創見解。

陳昇瑋在清華大學資訊科學系就讀期間，早已經開始投入很多心力在 IT 上，學生時代即鋒芒漸露。陳昇瑋最想要做的是具有獨創和原創特性的事情，一旦達到他所預期的成效，會讓他更清楚地知道「沒有我，就沒有這個研究」，因為挑戰無窮，所以覺得每天都過得非常充實，而且明白自己的價值。

所以他選擇了學術這條路，成為學者，他不光研究理論，也繼續寫程式。他覺得，將兩件事做到 85 分很厲害，但將一件事做到 95 分更重要。從程式人到投身學術界，作家劉墉說過的一段話影響陳昇瑋做了這個決定。這段話的意思是如果你今天在一個山頭上，看見另一個山頭想上去，第一件事得先下山，因為你不可能同時站在兩個山頭上，這是劉墉對於自己從中視記者毅然離開、成為作家的告白，對陳昇瑋也很感同身受，他只想完全歸零，連以前 BBS 的 ID（Xshadow）都不再使用。

陳博士期待，網路系統將會更聰明，能更有效的回應使用者的需求。

更詳細的介紹請參照 page10-1,10-2。

有人訪問 H.J. Eysenck（一位知名的心理學家）：「當某個研究帶來令您滿意的成果時，您的感覺如何？」

他的回答是：「那種感覺就好像貓得到奶油一樣，很難加以形容。我想那是一種充盈的美好感覺，你覺得一切都是那麼地令人愉快。我個人對自己原先預測的某些東西終於成為事實，多少會感到驚訝，因為我總覺得那不太可能，但現在卻成為事實，實在值得慶祝。所以你會對自己感到高興，對自然、對整個世界感到高興，你覺得生命真美好，值得繼續活下去。」

你也想一起來體會「貓得到奶油」的感覺嗎？

陳昇瑋博士愛貓，這是他的貓，可愛的毛毛跟小丁。

這就是陳昇瑋博士給所有對學術研究有興趣者的公開邀請。

從程式人跨界學術人的深度歷險

採訪 / 李宗翰 ・ 撰文 / 劉怡麟

中央研究院資訊科學研究所助研究員陳昇瑋，放下過去的名聲、只帶著多年累積的開發能力，一心勇闖學術界，以便更自在地鑽研開發所喜好的技術主題，追求無可取代的自我價值。

● 學經歷：國立臺灣大學電機工程博士，國立清華大學資訊工程系碩士／學士。原名陳寬達，曾擔任 Windows 及系統程式設計師、業界程式設計課程講師、技術書籍及雜誌專欄作者、共享軟體作者及 BBS 開發者／站長等工作，著有《Delphi 深度歷險》、《C++Builder 深度歷險》等書。
● 個人網址：http://www.iis.sinica.edu.tw/~swc/index_c.html

Delphi、C++Builder 過去曾是許多人撰寫 Windows 程式碼的首選，而推廣這兩套開發工具的先鋒人物中，有一位特別引人注目。他是陳昇瑋（原名陳寬達），《Delphi 深度歷險》與《C++Builder 深度歷險》的作者，於 1995 年與同樣熱中程式開發的錢達智，合力創辦 DelphiChat/BCBChat 電子郵件討論群（即 Programmer 深度論壇網站的前身）。

學生時代即嶄露頭角

早在陳昇瑋就讀於清華大學資訊工程系時，便已在 IT 界闖出名號，即使是對程式一竅不通的大眾電腦使用者，對他也不會陌生。他曾擔任紅極一時的楓橋驛站 BBS 站長（代號 Xshadow）、創立 Delphi 與 C++Builder 的程式設計技術討論社群，並擔任《HopeNet 光碟月刊》技術編輯，為讀者精選許多好用的共享與免費軟體。他自己也寫出不少廣受好評的共享及免費軟體，包括 XDesktop、XTheme Manager、XMP3 Manager、XViewer、XChat 等。

儘管陳昇瑋在 IT 界起步甚早，卻在 2000 年以後逐漸淡出。他放棄成為高薪 IT 人的機會，投入薪資相對較少，但自由度相對較高的學術領域。他目前身分為中央研究院資訊科學研究所暨資訊科技創新研究中心助研究員，並主持多媒體網路與系統實驗室。他對閱讀論文及發表研究成果樂此不疲，玩心重的他將論文投稿視為一場與全世界研究者的競賽，覺得「非常有挑戰性、有趣，且很刺激。」

為理想毅然轉身，尋找無可替代的個人價值

之所以離開 IT 界、投入學術圈，陳昇瑋提到下列幾個原因：

首先，他很早就體悟到，進入 IT 業界，不可能只管 IT 技術。再怎麼一等一的高手，時間體力都會被綿綿無絕期的會議和客戶問題消磨殆盡。

陳昇瑋大四時，曾考慮以 SOHO 方式講課和接案維生，如此在工作之餘，可以按自己的步調做研究或進修。他先後在中科院、新竹科學園區、資策會教育訓練處等講授程式開發課程，收入頗豐，卻漸漸發現，這條路無法如願。他觀察到幾位成功的前輩，並非想像中無拘無束，依舊得花大量時間自我行銷和經營。他也有過創業的想法，於 2000 年與朋友合夥創立軟體公司，但他發現要取得市場上的成功，不單純是技術問題，更多的是非技術上的挑戰，鎮日為資金、時間、人力及行銷問題忙得焦頭爛額，無法定下心來思考。

他說：「在學術界，你比較有時間冷靜客觀地去看世界面臨的問題；做的事情或許相同，但可以更有餘裕，不用擔心市場、時間，不用擔心公司裁員。」因此他毅然轉身，時至今日，他沒有後悔過自己做這個決定。

再者，業界走向取決於市場需求，總是

跟著流行走，而「財力」在這場混戰中扮演關鍵角色。陳昇瑋對市場的觀察是：有些好的想法沒有足夠財力支持，便不可能實現；同件產品由大小公司開發，最後往往是財力大或行銷實力強者勝出。「在這點上，相較之下學術界還是平等些，」他說。「一個十人團隊，不見得就會贏過三人團隊，重點在你的想法是否夠新穎、夠具獨創性。」

此外，業界的績效認定無可避免地必須是「能否幫公司賺錢」，但其中變數包括市場、時機（天時地利人和）甚至大環境等，絕非一己力可控制。相較之下，在學界「努力通常會有一定的回饋」。只要論文投上好的會議，好的期刊，付出不會被抹滅。

而最後也最關鍵的考量是，陳昇瑋理解到在很多工作當中，「你」很重要，但沒有「你」，還是能做到。

陳昇瑋認識不少投入業界的朋友，他們喜歡自己的工作，也有一定的成就感。然而，公司的陣容愈龐大，就愈難彰顯個人的重要。他讀過某退休副總裁寫的一篇文章，內容提到，不管你的位階再高，待遇再好，若有天離開（自願或非自願），公司照樣可以運行，你的工作會有人接手繼續，個人價值很難看得見。

但在學術界，你不會被逼著解決某個既存問題，相反地，你可能提出一個不曾被發現的問題，而這個研究，「沒有你，就不會存在，」陳昇瑋說。在浩瀚的學術領域裡，值得研究的議題多如繁星，絕對不用擔心做完，且完全可由自己主導。「業界的人不見得做不到，」他說，「但是他們沒有辦法五年十年都專注解決一個基礎性的問題。」

因為挑戰無窮，陳昇瑋每天都過得非常踏實、充滿自信，且明白自己的價值。

成為學者，他不光研究理論，也捲起袖子寫程式

身為資深程式設計師，陳昇瑋發現自己在學術界的優勢。他研究電腦及網路系統效能表現，如 Skype、MSN Messenger 及 AIM 等即時通訊軟體的語音及視訊品質，為量測及監看網路及軟體的行為及效能，經常需要寫程式。「像這類程式，程式碼的長期維護性不是最重要的，但一定要寫得夠快、夠有彈性。」在研究試誤階段，有太多因素不可預測，一個程式改個十幾二十次是家常便飯，因此程式開發功力和效率十分重要。

實務上，他們不特別關注特定平臺的解決方案，跨平臺的開發是家常便飯。以視訊串流的實驗來說，若有某系統提供好用的函式庫，他們就會想辦法搭配使用。因此有可能網路頻寬以 FreeBSD 架設的路由器來控制，而視訊串流的伺服端及客用端就用 Microsoft Windows。對他來說，一個實力堅強的程式設計者，不會受限於任何開發工具及平臺，因為他已經掌握了關鍵概念及技術，且懂得常用資料結構及演算法的特性及應用時機，常用的設計模式（design patterns）也都在腦中，因此若需要用到動態規劃（dynamic programming）或優先權佇列（priority queue），隨時可以寫得出來，而不只是會用直覺及暴力法來解決開發需求。

> 在學術界，你可能提出一個不曾被發現的問題，而這個研究，「沒有你，就不會存在」陳昇瑋說。

陳昇瑋近幾年較少專注在程式開發，但遇上不熟悉的環境或語法，依然可以快速上手；學生撰寫的程式品質優劣，也能輕易看出。他提到自己實驗室的學生經常困在程式問題，以致實驗進度延宕。有一次，他請學生寫一支程式，給他兩週時間。兩週後，學生徒勞而返，他失望難免，於是將學生拉到會議室，在他面前邊解說邊撰寫，一個字一個字敲，半小時後寫就。這件事讓那位學生印象深刻，一來少有老師會親自示範寫程式碼，二來是他兩個禮拜都做不出來的東西，老師只需要半小時。

從 IT 人跨界為學術人，陳昇瑋的夢想沒有變。他帶著累積十數年的程式開發實力，到學術界來做技術創新，將理論與實務做最完美的結合。

將兩件事做到 85 分很厲害，但將一件事做到 95 分更有價值

與陳昇瑋背景相似的 IT 人中，很少有人跟他做同樣的選擇。作家劉墉說過的一段話影響了他：「如果你站在這個山頭，羨慕另一個山頭更美，第一件事，就是走下這個山頭！」當時劉墉是一位高薪主播，為了理想毅然辭職，追尋自己的畫家夢。當陳昇瑋決定成為學術人時，心裡浮現的正是劉墉的這番話，因為一個人不可能同時站在兩座山頭上。他決定讓一切歸零，重新出發，就連過去的 BBS 帳號 Xshadow 都拋棄。

對於人生方向，陳昇瑋給年輕人的建議是：瞭解自己（“Know thyself.”），想清楚再決定，但確定後就要心無旁騖，才能將它做好。他所謂的「好」指的是登峰造極。他說，能將兩件事做到 85 分，那表示很厲害，但將一件事做到 95 分更有價值。

全心全意、專注心思很重要。陳昇瑋在學生時代，經常有講課及接案的機會，賺小錢易如反掌，最好的個人記錄是在 3 天內完成 20 萬元的案子。但考上博士班後，他決心婉拒所有邀約，儘管他的同儕認為利用週休或假日賺點外快沒什麼大不了，他卻不這麼想：「原本該休息的時間用來賺錢，就勢必難免會排擠工作時間。」他也幾乎不漫無目的地看電視或報紙，改以閱讀趨勢專文或專書的方式保持對社會的關注：「真正重要的資訊，你不會不知道。不要讓資訊焦慮影響你，那只會讓自己被資訊洪流淹沒」。

他說：「如果你有空跟別人說你很忙，如果你還記得打開 BBS 或看新聞網站，那表示你還不夠忙、不夠投入。」他並非要大家一年 365 天都投入工作或學習，但他覺得有潛力就該盡量發揮，向極限挑戰，看看自己能做到哪裡；除了大量閱讀，更要動手做（觀察、計算、寫程式），這樣，創意才能不斷撞擊，無限蔓生。

再看二元搜尋法

• 穆信成 •

如果你自認對二元搜尋夠熟悉，卻沒讀過 N. van Gasteren 和 W. Feijen 的筆記 The Binary Search Revisited, 強烈建議你找時間看看。

資訊系學生一定學過二元搜尋：給一個排序好、長度為 N 的陣列，在 $O(\log N)$ 的時間內決定其中是否有某個值。Bentley 在 Programming Pearls 一書中卻提到只有大約一成的專業程式員能第一次把這程式寫對。這一定要親身試試看才會相信。

迴圈、不變量、與界限

Van Gasteren 和 Feijen 在一篇 1995 年的研究筆記中釐清了一個迷思：您可知道，二元搜尋其實並不一定要用在排序好的陣列上？事實上，他們認為把二元搜尋類比為翻字典找字，反倒是個盲點。他們先試解一個更廣的問題：給定整數 M 和 N, 和接受兩個整數參數的布林函數 Φ，已知 $M < N$, 並知 $\Phi(M, N)$ 成立。Φ 另需滿足的條件待會兒再談。試找出兩個介於 M 和 N 之間、滿足 Φ 的相鄰整數。若寫成邏輯式子，程式執行完畢後需滿足：

$$M \le l < N \wedge \Phi(l, l+1)$$

這是他們的程式：

$\{M < N \wedge \Phi(M, N)\}$
$l, r := M, N$
$\{\text{Inv}: M \le l < r \le N \wedge \Phi(l, r),\ \text{Bound}: r - l\}$
; do $l + 1 \ne r \rightarrow$
  $\{l + 2 \le r\}$
  $m := (l + r)/2$
 ; if $\Phi(m, r) \rightarrow l := m$
  [] $\Phi(l, m) \rightarrow r := m$
  fi
 od
$\{M \le l < N \wedge \Phi(l, l+1)\}$

虛擬碼中紅色字是註解，也是斷言 (assertion)，表示程式執行到此處一定會成立的條件。

如何得知本程式正確呢？讀者不妨試著檢查看看：

- 標為 Inv 的「不變量 (loop invariant)」$M \le l < r \le N \wedge \Phi(l, r)$ 在進入迴圈前成立。$l, r := M, N$;
- 若迴圈執行前 Inv 成立，迴圈執行一次後 Inv 仍成立。
- 標為 Bound 的「界限」$r - l$ 最初是正整數。在迴圈中 m 的值必在 l 和 r 之間。迴圈每執行一次，$r - l$ 就變小。當其等於一時，迴圈就得停了。
- 迴圈停下時，l 和 r 已是相鄰的整數。

和 Inv 合起來看，我們要的結束條件 $M \le l < N \wedge \Phi(l, l+1)$, 就滿足了！

> 在一篇 1995 年的研究筆記中釐清了一個迷思：二元搜尋其實並不一定要用在排序好的陣列上。

只是還有一個但書：if 敘述中的 $\Phi(m, r)$ 和 $\Phi(l, m)$ 這兩個條件至少要有一個成立。這是 Φ 的另一個要求：

$$\Phi(l, r) \wedge l < m < r \Rightarrow \Phi(l, m) \vee \Phi(m, r) \quad (*)$$

有哪些函數 Φ 滿足條件 (*) 呢？Van Gasteren 與 Feijen 舉的例子包括：

- $\Phi(i, j) = a[i] \ne a[j]$ 此處 $a[M...N]$ 是某陣列。程式會幫我們找到兩個相鄰但不相等的元素。Van Gasteren 與 Feijen 認為這個特例可能是解說二元搜尋更好的例子。
- $\Phi(i, j) = a[i] \times a[j] \le 0$; 找兩個相鄰但正負號不等或至少有一個為零的元素。
- $\Phi(i, j) = a[i] < a[j]$,
- $\Phi(i, j) = a[i] \wedge a[j]$, 等等。

Van Gasteren and Feijen 指出了令人驚訝的事實 ...

搜尋已排序陣列

回到最初的問題：如何在排序好的陣列中找某個關鍵值呢？理想上我們希望設計某個 Φ 並套用前面的程式。第一個念頭是令 $M = 0$ 並用這個 Φ：

$$\Phi(i, j) = a[i] \le K < a[j]$$

演算法結束後 $a[l] \le K < a[l+1]$ 必成立，只要再看看 $a[l]$ 是否等於 K 就可以了。但問題來了：$\Phi(0, N)$ 最初得要滿足，但 $a[0] \le K < a[N]$ 並不見得成立呢。

Van Gasteren 與 Feijen 建議在陣列頭尾各補一個想像元素：$a[-1]$ 無限小 $a[N]$ 無限大。等價的說法是用這個 Φ：

$$\Phi(i, j) = (i = -1 \vee a[i] \le K) \wedge (K < a[j] \vee j = N)$$

最初令 $l, r := -1, N$, 那麼初始條件就滿足了。如果在陣列頭尾補元素讓你覺得很不妥，別擔心。Inv 保證 $-1 < m < N$，因此從頭到尾，$a[-1]$ 和 $a[N]$ 兩個位置都沒被讀過。陣列並非真有改變，兩個想像元素只是為了證明演算法正確而假想的。

有興趣的讀者可練習看看：有個陣列先遞增、後遞減，如何用二元搜尋找到頂點？

其他

一般的二元搜尋程式在找到關鍵值時會用 break 跳出迴圈。這會比較快嗎？為何前幾年又有人說世界上大半的二元搜尋程式還有 bug ？把 m 設為 $(l + r) / 2$ 好不好？

欲知更詳細的內容，請看筆者的部落格！ http://www.iis.sinica.edu.tw/~scm/ncs

The Institute of Information Science (IIS) at Academia Sinica, Taiwan, R.O.C. invites all qualified candidates to apply for the positions of junior and senior research fellows of all ranks (equivalent to the ranks of tenure track assistant, associate and full professors in a regular academic department without teaching responsibility) in all areas of Computer Science. In particular, candidates in the areas of computer systems, machine learning and natural language processing are strongly encouraged to apply.

Academia Sinica is a national academic research institution in Taiwan that conducts research on a broad spectrum of subjects in science and humanities. IIS is committed to high quality research in computer and information science and engineering. In addition to research funding supported by Academia Sinica, external funding through government agencies and industry-sponsored institutions is also available.

Full-time research fellows are free to set their own research directions. IIS currently has about 40 full-time research fellows and close to 300 full-time post doctoral fellows and research assistants. The areas of their current research include Systems Technology, Bioinformatics, Multimedia, Natural Language and Knowledge Processing, Network and Theoretical Computer Science and Parallel Processing.

All Candidates should have a Ph.D. degree in computer science or closely related fields with a strong research and publication record. Senior candidates must demonstrate strong leadership, and have an international reputation evidenced by publications, patents, industrial experiences, or other academic and scholarly achievements. Salary is commensurate with qualifications.

All candidates should send detailed curriculum vitae, and at least three letters of recommendation to

Dr. Pen-Chung Yew, Director
Institute of Information Science
Academia Sinica
Nankang 115, Taipei, Taiwan
Tel: 886-2-2788-3799 ext.1729
Fax: 886-2-2782-4814
E-mail: recruit@iis.sinica.edu.tw

Fluency in Chinese is an advantage, but not required. For additional information about IIS, please visit http://www.iis.sinica.edu.tw.

Institute of Information Science
Academia Sinica
128 Academia Road, Section 2, Nankang, Taipei 115, Taiwan, R.O.C.
Tel: 02-27883799
http://www.iis.sinica.edu.tw/

Publisher: Director Pen-Chung Yew
Editors: Pei-Chi Wang and Huey-Chyi Chris Tseng