

生物網路的定量模擬方法研究

蒲泰宏

國立暨南國際大學

資訊管理研究所

s94213517@ncnu.edu.tw

俞旭昇

國立暨南國際大學

資訊管理系

ssyu@ncnu.edu.tw

蔡劍霞

國家高速網路與計算中心

jstsai@nchc.org.tw

摘要

在基因體定序完成後，了解生命體內的運作與現象等重要工作即將肇始。生命現象的複雜，並非肉眼可及，更遑論可以用一般的科學方法去觀察或解釋。在系統生物學的機制下，透過 dry-lab 和 wet-lab 的相互支援與配合，可以幫助我們探究這瑰麗的寶藏。

本研究擬建構一套模擬系統，初探生命網路中元件間互動的現象。別於過往的經驗，我們捨棄複雜的數學模型，透過網路的概念，去描述一個生命交互作用的情況；更透過修正的 Petri net，加入了物件導向技術的觀念，使整體模型的描述更能貼近真實現象。

經由此項研究，可以幫助我們對於藥物發展以及病例研究有相當的幫助。在許多實驗是無法藉由人體直接實驗或觀察，透過這樣的模擬系統，不僅可以節省開發藥物的時間，更可以免除不必要的實驗風險。

目前本研究以醣酵解作用(glycolysis)以及檸檬酸循環(TCA cycle)作為測試標的，透過參考文獻修正系統的不足，期許能應用於更複雜的生物網路，如細胞循環、訊號傳導或蛋白質交互作用網路等。

關鍵詞：系統生物學、生化網路、Petri net、物件導向。

生物網路的定量模擬方法研究

摘要

在基因體定序完成後，了解生命體內的運作與現象等重要工作即將肇始。生命現象的複雜，並非肉眼可及，更遑論以一般的科學方法去觀察或解釋。在系統生物學的機制下，透過 dry-lab 和 wet-lab 的相互支援與配合，可以幫助我們探究這瑰麗的寶藏。

本研究擬建構一套模擬系統，初探生命網路中元件間互動的現象。別於過往的經驗，我們捨棄複雜的數學模型，透過網路的概念，去描述一個生命交互作用的情況；更透過修正的 Petri net，加入了物件導向技術的觀念，使整體模型的描述更能貼近真實現象。

經由此項研究，可以幫助我們對於藥物發展以及病例研究有相當的幫助。在許多實驗是無法藉由人體直接實驗或觀察，透過這樣的模擬系統，不僅可以節省開發藥物的時間，更可以免除不必要的實驗風險。

目前本研究著以醱酵解作用(glycolysis)以及檸檬酸循環(TCA cycle)作為測試標的，透過文獻的參考，將修正系統的不足，並期許更能廣泛應用於其他領域如細胞循環、訊號傳導或蛋白質交互作用網路等研究。

關鍵詞：系統生物學、生化網路、Petri net、物件導向。

1. 前言

生物學早期發展就像拼圖一樣，經過不同的觀察描述所看到的現象，企圖藉由將這些已知的片段組成生命的輪廓。傳統上以化約論看待生命活動，視生命體內的各個子系統都能獨立運作，現在的生物技術發展卻印證子系統之間存在互動關係，進而衍伸複雜的生命行為。

但是推論或者模擬是否能真正準確的預測出

真實的生命現象？這是一大難題。化學反應以化學方程式做定性描述，再轉換成數學模型進行定量。代謝工程(Metabolic Engineering)往往透過微分方程組(ODEs)建模才能定量求解。但數學模型有其計算上的複雜與延展性不佳之缺陷。從數學模型著手通常有三個方法如表一：[15]

表一 建模方法與挑戰

方法	挑戰
線性函數	無法準確描述現象 與穩定的震盪
分段線性	片段如何分割 和段點的處理
非線性現象	單靠數據驗證 是否代表模型一定正確

資料來源：本研究之整理

由上述可知，數學模型在生物意義上的有效性是一大課題！此外，建模方式造成計算的複雜度隨系統元件數量劇增，且模型的彈性不足以因應環境變動，需大幅修改。計算上的誤差是否在容許範圍內，也是影響實用性的關鍵。利用資訊管理的觀點，我們將生物環境圖形化，以資料結構(Data Structure)中圖形(graph)的概念，將所有生命元件視為網路組態。圖形化易於直觀生命活動，將各個參與份子如反應物、ATP、H₂O、視為所要傳遞的元件，而發生反應的過程則視為節點，透過節點之間的連結，構成一個無尺度的網路(scale free network)。且研究發現，生物網路具有小世界網路(small world)的特性，具有高度的容錯能力。因此這樣的資料結構可以幫助我們建構定性模型(qualitative model)，再使用 Petri net 來幫助完成定量模型(Qantitative model)的建構。Petri net 不僅可以呈現靜態地元件間關係，由於具有數學上的驗證，

更可合理呈現元件間動態行為。

在計算上利用多執行緒的技術模擬整體反應的進行。反應並非循序發生，而是一種離散事件 (Discrete Event) 的條件觸發，例如溫度、壓力、基質的濃度或酵素的參與而調控整個反應的進行。

在生命體內不同的反應空間，如胞膜或胞器有隔開反應的作用，在本研究稱為 Compartment，於開發系統時引進物件導向 (Object-Oriented) 中容器 (Composite) 的概念，作為不同生命體中的反應容器。

在開發系統方面，我們有別於目前的模擬軟體，目標在設計出能適用於不同領域的模擬架構，而不單只為了模擬代謝網路或者生物相關的案例。透過另外設計好的演算法及搭配所需參數，將可以套用於不同領域的需求與研究。

2. 文獻探討

根據前述，本研究利用 Petri Nets 模型建構醣酵解作用以及檸檬酸循環代謝網路模型，以特定的資料結構和修正後的動力學計算模型作整體網路動態模擬，因此本節將介紹模擬的方法、Petri Nets 以及代謝網路。

2.1 模擬方法

系統生物學是整合不同領域所產生的新興研究面向，促使生物學走向預測反應結果，而電腦模擬是很適合的途徑。在這前提下，須先收集許多各種生物反應的數據，如基因體、蛋白質體以及代謝體等，透過建立計算模型，將最終的結果與其相比，評估模型是否需要修正[4][5]。

模擬的好處，由以下幾點作說明[6]：

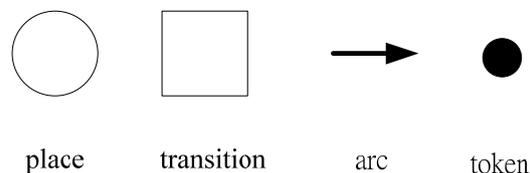
1. 可有效預測或推論實驗數據。
2. 在實驗環境不便時，可作為協助實驗的工具。
3. 節省金錢與時間成本。
4. 以客觀方式了解生命系統的流程。
5. 能探究隱含的資訊。
6. 透過視覺化更能幫助理解複雜的生命流程。
7. 驗證與支持一個生命系統所產出的定量資料。

誠如上述，在開發軟體時必須對模擬的方法做

評估與選擇。模擬的類型通常可以分為與時間是否有關？如固定性 (static) 模擬和動態性 (dynamic) 模擬；與模擬次數有關？如確定性 (deterministic) 模擬和隨機性 (stochastic)；或者是具有連續時間發生變化的連續型 (continuous) 模擬或者是由事件所驅動的，紀錄不同時間點所呈現之不同狀態的離散事件 (discrete event) 模擬。在研究[6]中，作者對各類模擬的方法做整理與描述，傳統以微分方程組 (ODEs) 為開發依據，用數值方法建立一連串的反应速率方程式，再輔以圖表觀察反應物質隨著時間變化而增減的情形。這種與時間相依 (time dependent) 的方法，雖然具有堅實的數學學理作後盾，但是其建立和修改方式困難是一大弱點。其它如 PDE、Stochastic differential equations 以及由 non-linear ODEs 所簡化的 S-system 等方法，亦無法逃離數學計算上的繁雜之苦。而 Petri Nets 的方法正好彌補缺憾，不但使建模簡單化，且同時表達定性與定量的關係，不僅有數學驗證的基礎，也有視覺化的呈現，很適宜用在生化網路研究。故本研究以 Petri Nets 為開發工具，並於後簡介 Petri Nets 的符號與定義。

2.2 Petri Nets

Petri net 是 C.A. Petri 於 1962 年的博士論文所提出[9]，再由 MIT 等研究人員的研究，使其完善成為一套理論基礎，它不僅具有圖形化且具有數學化的正規模擬工具[1]。在圖形的表示方面如圖一所示：



圖一 Petri net 組成元件

資料來源：Petri (1962)

基本的模擬網路是由此四個元件所構成，place 代表輸入或輸出的位置，通常用來表示狀態 (condition)，可在其中置入 token；而 transition 表示事件 (event)，兩者會透過 arc 連結。當滿足觸

發條件(如 token 數足夠或者大於 arc 的權重等)則 transition 觸發,使 place 的狀態改變[10]。之後的研究者亦將 Petri net 變形,稱為 high-level net,如 Hierarchical Petri nets、Hybrid Petri nets、Timed Petri nets 等,以符合更複雜的模擬需求。圖形化的魅力正是傳統技術無法望其項背的原因,傳統上利用數學如微分方程所開發的系統(如 Gepasi、E-Cell 等)或是以 Pi-Calculus、Bio-Calculus 等架構的系統,面臨的問題即是無法輕易編輯生化路徑[7],而數學公式的繁雜,更容易造成計算上的錯誤以及非相關領域的人無法輕易了解各個參數或函數,在生物現象的對應以及意義。此外, Petri net 具有(1)可讀性(readability)–有助於了解並避免建構模型時的失敗,(2)可行性(executability)–此實驗的模型能夠讓人輕易了解,(3)可驗證(validation techniques)–確保模型的完整性與符合真實性,(4)可分析(analysis techniques)–可做定性及定量模型的行為預測[3]等特性,對系統生物學研究有很大助益;甚者, Petri net 的變形也適用於不同的生化路徑,例如文獻[4]提到代謝物如何在生化路徑中產出,可以將相關的基質、產物以及抑制因子等量化後計算;[5]抑制因子可能會出現在代謝網路中;[8]生物路徑會發生同步處理等,都明白表示, Petri net 有足夠的能力以及理由來建構一個生物網路,並且表達這個網路的處理流程 [3]。

Time Petri Net (TPN) 是典型的 Petri Net,由 $[a_t, b_t]$ 代表某一 transition t 可能發生反應的時間區間,其中 a_t 和 b_t 為自然數(可包含 0),且 $a_t \leq b_t$ 或者 $b_t = \infty$ [8][12]。其正式定義如下:

$Z = (P, T, F, V, m_0, I)$ 為 Time Petri Net, 若且唯若

(1) $S(Z) = (P, T, F, V, m_0)$ 是一個 Petri Net,

(2) $I: T \rightarrow Q_0^+ \times (Q_0^+ \cup \{\infty\})$ 且 $I_1(t) \leq I_2(t), \quad t \in T, I(t) = (I_1(t), I_2(t))$ 。

則可以稱 TPN 是一個 finite Time Petri Net (FTPN)

若且唯若 $I: T \rightarrow Q_0^+ \times Q_0^+$ 。

這裡 Z 結構內有六個要素, P 代表一個有限個數 Place 的集合; T 代表一個有限個數 Transition 的集合; F 代表一個射線的集合; V 代表一個權重方程式; m_0 代表初始標記; 以及最後的 I , 代表時間

函數。透過 TPN 此模型,將可以用於需要考慮時間因素的模擬環境,本研究所使用的模型類似 TPN 的概念,即是必須考量代謝路徑中各化學反應之反應時間有所差異,必須做排程處理。

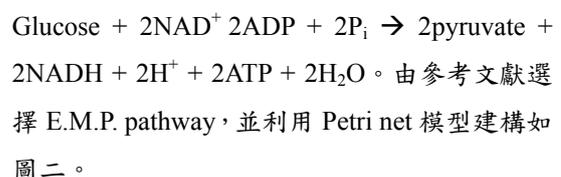
此外,在我們所使用的模型中,運用了物件導向(Object-Oriented)中,平行處理(parallel processing)的概念。傳統的 Petri net 對於路徑選擇有其局限性,對於物質該如何決定其方向有設計上得缺陷,因此藉由平行處理,將可以更貼切真實現象,反應並非只有循序發生,且無存在絕對的單一路徑傳遞物質的準則,因此透過平行處理的方法,將可同時大量處理的反應情形。

2.3 代謝網路

本研究利用 Petri Net 模型分別建構下列兩者代謝網路,分別為:

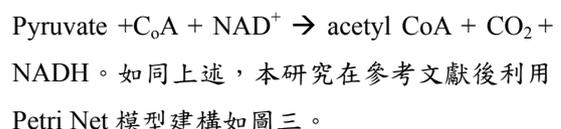
1. 醣酵解作用(glycolysis):

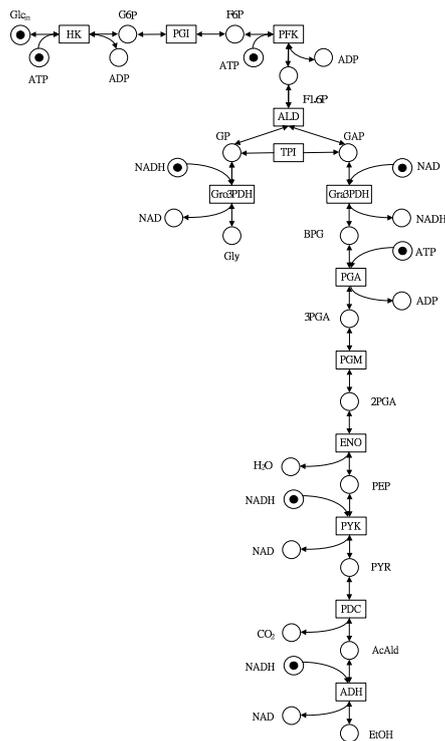
醣酵解作用是指將葡萄糖(glucose)轉變成丙酮酸(pyruvate)的過程,其淨反應式如下:



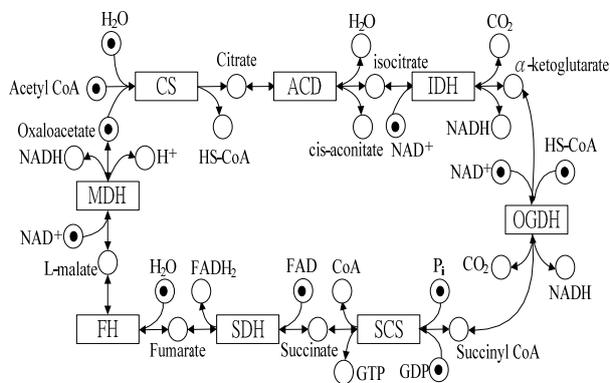
2. 檸檬酸循環(TCA cycle):

檸檬酸循環是假設醣酵解作用在持續有氧(Aerobic condition)下,所產生的丙酮酸會由細胞質液進入粒線體,開始開始檸檬酸循環過程。會使丙酮酸完全氧化成 CO_2 和 H_2O , 並將葡萄糖內大部分的能量釋放出來。其淨反應如下:





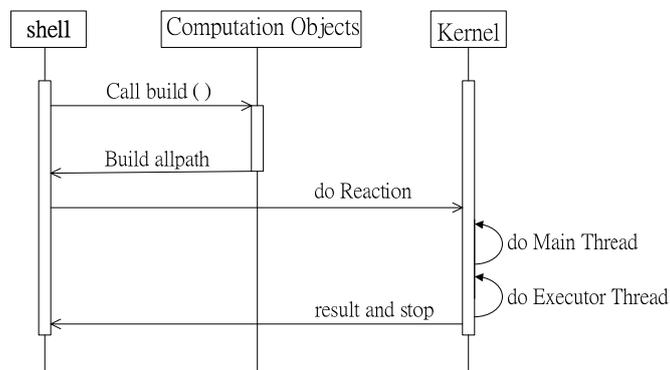
圖二 糖酵解作用
資料來源：本研究之整理



圖三 檸檬酸循環
資料來源：本研究之整理

尋適合的模型來描述現象；接著尋找目前可獲得的資源與工具(例如模擬的系統或工具)；然後要安裝系統並學習如何操作...等。以上的前置作業完成後，仍未開始模擬實驗[13]。雖然有研究者提出用網站提供模擬服務[14]，但是對於建模的方式並不彈性。在系統生物學的觀點下，每一個 reaction 為一個研究的單元，透過許多單元的組成來研究彼此互動，且實際的生化網路中，物質的傳遞或走訪並非唯一路徑。因此，若將路徑的樣板寫死(即使用者無法修改或新增、刪除節點等)並不是最佳的模擬方法。故在建構代謝網路時，我們以 reaction 為研究單元，建立相關物質之參數，並依照不同的研究設計、實驗觀點對路徑作選擇性的修改。

在系統架構上，可視為以下三者間的互動：前端為 Shell，為各種與計算物件媒合的模型格式；中間層為 Computation Object，是模擬計算的程式；底層為 Kernel，系統核心用以管理資源與分派資源，如圖四：



圖四 系統架構
資料來源：本研究

3. 研究設計

本研究目標建立有別於其他軟體的模擬系統，不將系統侷限於代謝網路模擬，更希望能進行多元化研究，如基因調控網路、訊號傳導網路，或是其他領域的模擬研究等應用。

3.1 系統架構

在過去的經驗裡，執行模擬要先釐清問題，找

使用者透過 Shell 呼叫函式 build 建立欲研究的路徑模型，因此可以找出網路中所有待執行模擬的反應節點。這些反應節點，Computation Objects，會依照執行時間的不同做排程。因此，在 Kernel 中，負責資源管理及分派的工作，主要執行兩種 Thread：Main Thread 和 Executor Thread。除了 Executor Thread 負責將等待執行的反應從佇列取出執行，直到佇列為空之外，Main Thread 負責做資源分配以及反應的排程。

1. Kernel

系統的底層由 Class Simulate 負責，其功能為系統資源的分配與管理，這裡指稱的資源為要參與計算的相關反應物。因此 Simulate 並不涉及相關的領域知識，此設計可以保持系統的彈性，使其可以媒合不同的領域作相關研究。Simulate 可從檔案或者資料庫中讀取要計算的反應物並給予其編號，透過這個識別碼作資源的控管以及演算法做排程處理。

2. Computation Object

目前著手於代謝網路的模擬，故以網路中的 reaction 為模擬單元，或稱之為「計算物件」，此利用化學平衡、熱力學以及酵素動力學 (Michaelis-Menten rate law) 等概念所勾勒出的計算方程式。在計算方面，我們引用[2]的修正方法，其將熱力學的參數結合酵素動力學後，只需將相關酵素特定值儲存於資料庫，經過呼叫後即可進行計算。計算物件依據該 reaction 取得或釋放資源、檢查是否會觸發其他 reaction、計算反應完成的時間和計算反應是否達到動態平衡等。目前時間的計算，依照酵素動力學的推導，反應是否向右進行是根據過渡時期轉變成產物時的濃度而定，稱為 K_{cat} ，又稱為轉化數 (turnover number)。其倒數 (即 $1/K_{cat}$) 為反應時間，是一常數。未來可能假設環境會受到其他物質的干擾，如抑制劑 (inhibitors) 或者活化劑 (activators) 參與反應，應將目前以常數代表反應時間更改為一段計算式，透過即時運算求出反應的時間。

計算效能的提昇方面，除了電腦強大的運算能力之外，以酵素動力學透過微分方程的計算，牽涉到時間切割的問題，這不僅牽涉到時間單位要切割至多細微才具有精確性之外，初始濃度的設定以及操作點的選擇，往往會造成參數估計陷入局部性的窘境，透過斜率逼近是否準確是一個難題外，更重要的是，偏微分方程在計算上往往耗費時日！本文中的方法，為透過二分逼近法求得最終的平衡解，我們著眼

於最終的結果，但並非刻意忽略中間的細部過程，畢竟生命體內的反應，以及物質傳輸的速率，其速度之快實在難以估計，況且假設於臨床實驗所需為最終的結果參考，而研究的節點數目並非以數百計，要同時觀察大量的反應節點，若透過微分方程的建構，可想而知，勢必要消耗大量的時間與力氣，若模型無法重複使用，在成本皆是無情的負擔；透過目前的計算方法，則可節省大量的時間成本，在毫秒時間內即可完成整體模擬反應的計算。

計算物件除了目前代謝研究的應用外，尚可開發其他計算物件如應用於訊號傳導網路、基因調控網路或蛋白質交互作用網路等，或者其他非生物領域的應用，例如觀察網際網路中，網路的節點設備如同伺服器或路由器的流量研究等。

3. Shell

Shell 的部份，為模型的格式呈現。除了各個模擬軟體的檔案格式外(如純文字檔案)，亦可透過呼叫資料庫(如關聯式資料庫)建構。甚者，目前有標準化的趨勢，以 XML 格式為主的 SBML 或 CellML 等，皆為媒合各種模型格式的介面。Shell 的實作，除了滿足基本的模擬模型建構外，爾後更可以透過檔案格式交換的機制，將不同的模型格式轉換成統一的格式，置於網路上提供其他使用者參考，此優點不僅省略研究者重新建構模型的開發時間，更可達到跨平台使用的功效。

3.2 相關參數定義

系統所使用的計算物件，目前以修正後的酵素動力學為反應的計算方程式，因此在計算參數的引用，只需將其定義於資料表，爾後呼叫計算即可。於此，我們引用了相關參數有 ΔG_0 (自由焓)、 K_{cat} (轉化數)、 K_m (反應物反應常數)等酵素相關的常數值，並且透過運算求出 K (平衡常數)。相關參數可以透過 BRENDA 公開資料庫查詢，可由具有生化背景的研究人員將這些參數建置於資料表，而一般使用者未必了解這些參數的意義，所以在前面的呈現部

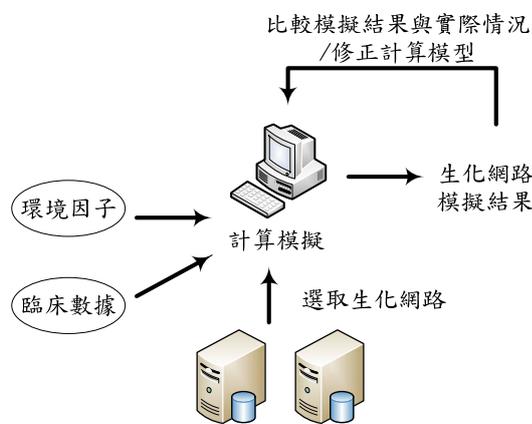
份，並不會涉及這些參數的相關資訊輸入，僅對欲觀察的反應物之濃度值輸入即可。

3.3 物件導向概念

在建構代謝網路時，為求更能貼近實際現象，我們引入物件導向 (Object-Oriented) 的觀念。細胞內的空間是非常擁擠且充滿著不同的物質如胞器、胞膜酵素及酵素複合體等。許多物質可能同時存在或者會透過通道彼此傳遞而運用。因此，對於細胞的結構，在設計模式 (Design Pattern) 中的 Composite Pattern 的概念，可說明這樣的現象，就如同許多大大小小的容器堆疊在一起。

3.4 系統佈署

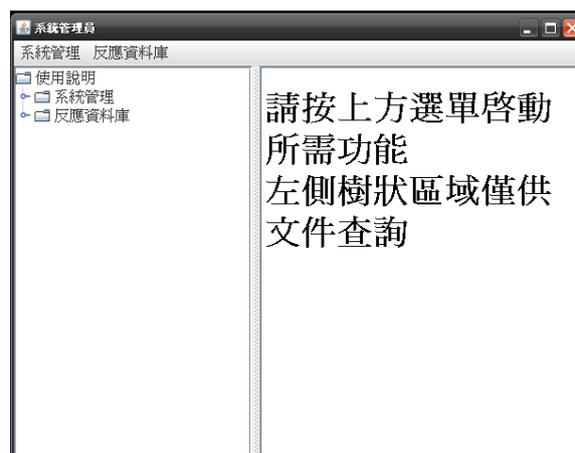
本系統利用 Java 語言開發，透過 Java 跨平台的特性，只要有安裝 JRE 的系統即可安裝並執行，提供使用者安裝與使用上的便利性。系統佈署如圖五所示，使用者可以透過事先定義好的規格，將其輸入並儲存於資料庫，日後可以讀取並執行模擬。在系統生物學的精神下，研究的方法如同中醫一樣，講求調和理論，即拜現代高科技所賜，能取得大量的豐富資料，透過這些資料執行模擬實驗，所得到的結果再與臨床實驗作比較，最後提出觀點說明結果或修正系統的缺失。本系統不僅提供使用上的便利，更希望能讓非相關背景的研究者也能夠輕易操作。



圖五 系統佈署
資料來源：本研究

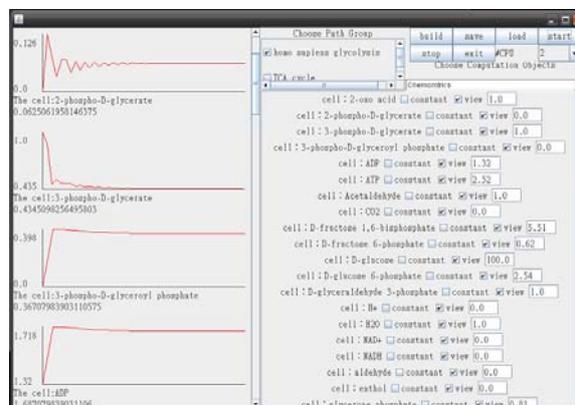
3.5 系統畫面簡介

在使用者登入系統後，可以看到系統畫面如圖六所示，其功能分為兩部份：系統管理與反應資料庫。系統管理為使用者帳號與密碼等相關資料管理；反應資料庫則對於反應常數資料、代謝路徑、反應資料等資料做管理，以及啟動模擬之功能。



圖六 系統管理介面
資料來源：本研究

圖七則為系統模擬畫面。在左半部分呈現反應物的濃度變化圖；在右半部分則可以勾選欲代謝路徑、相關計算物件選擇和基本的功能如存檔、讀檔、啟動模擬與停止等功能。在右下方的部份，則可輸入各個反應物的初始濃度以及勾選是否對其做觀察。



圖七 系統模擬畫面
資料來源：本研究

4. 結論與未來展望

透過電腦運算能力強大的魅力，可以幫助我們節省傳統利用數學模型計算所耗費的大量時間，並且提高計算的正確率。透過系統的模擬，將助於研究人員更能輕易的參與研究，只需依照事先定義好的資料結構與輸入規格，將數據依序輸入即可執行模擬，並且透過圖形化介面的結果呈現，更易於幫助研究人員解釋與了解所欲了解的標的。這樣的系統可以幫助任何不具深厚生化背景的人員也能輕鬆操作，執行系統模擬的研究。

目前研究為代謝路徑環境的模擬，利用醱酵解以及檸檬酸循環作為研究的對象，基於系統生物學的精神，在跨學科、跨領域的技術整合，在資訊科學的模擬結果出爐後，仍急需實驗數據的提供與驗證，由於臨床數據收集實屬不易，這是未來研究上需要加強的工作之一。

參考文獻

- [1] 李允中，劉豐瑞，「模糊派屈網路，模糊理論及其應用第 10 章」，全華出版社，2003/01。
- [2] Genrich, H., Küffner, R. and Voss, K., “Executable Petri net models for the analysis Of metabolic pathways”, *STTT*, Vol.3, 2001, pp. 394-404.
- [3] Heiner, M. and Koch, I., “Petri Net Based Model Validation in Systems Biology”, *ICATPN*, LNCS 3099, 2004, pp. 216-237.
- [4] Hood, L. , “Systems biology: integrating technology, biology, and computation”, *Mech. Ageing Dev*, Vol.124, 2003, pp. 9-16.
- [5] Ideker, T., Galitski, T. and Hood, L., “A new approach to decoding life: systems biology”, *Annu. Rev.Genomics Hum. Genet*, Vol.2, 2001, pp. 343-372.
- [6] Materi, W. and Wishart, D.S., “Computation al systems biology in drug discovery and development: methods and applications”, *Drug Discovery Today*, Vol.12, 2007, pp. 295-303.
- [7] Miyano, S. and Matsuno, H., “How to Model and Simulate Biological Pathways with Petri Nets – A New Challenge for Systems Biology”, *25th International Conference on Application and Theory of Petri Nets*, 2004.
- [8] Peleg, M., Rubin, D. and Altman, R.B., “Using Petri Net Tools to Study Properties and Dynamics of Biological Systems”, *Journal of the American Medical Informatics Association*, vol.12, No.2, 2005, pp. 181-199.
- [9] Petri, C.A., PhD Thesis, “Institut für Instrumentelle Mathematik”, Bonn, 1962.
- [10] Pinney, J.W. , Westhead, D.R. and McConkey, G.A., “Petri Net representations in systems biology”, *Biochemical Society*, Vol.31, No.6, 2003, pp. 1513-1515.
- [11] Popova-zeugmann, L., Heiner, M. and Koch,I., “Time Petri Nets for Modelling and Analysis of Biochemical Networks”, *Fundamenta Informaticae*, Vol.67, 2005, pp. 149-162.
- [12] oppova-zeugmann, L., and Schlatter, D., “Analyzing Paths in Time Petri Nets”, *Fundamenta Informaticae*, Vol.34, 2005, pp. 1-17.
- [13] Snoep, J.L., Olivier, B.G. and Westerhoff, H.V., “JWS Online Cellular Systems Modelling And The Silicon Cell”, *ESCEC^{5th} - 8th* , 2003.
- [14] Snoep, J.L., “The Silicon cell initiative: working towards a detailed kinetic description at the cellular level” *Current Option in Biotechnology*, Vol.16, 2005, pp. 336-343.
- [15] Torres, N.V. and Voit, E.O., “Pathway Analysis and Optimization in Metabolic Engineering”, *Cambridge University Press*, 2002.