

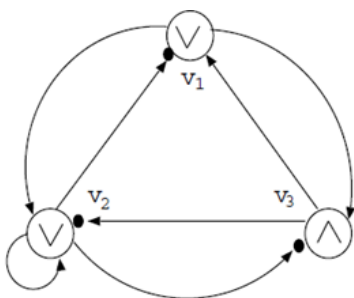
ブーリアンネットワークによる遺伝子制御ネットワークの解析

Analysis of gene regulatory network on Boolean model

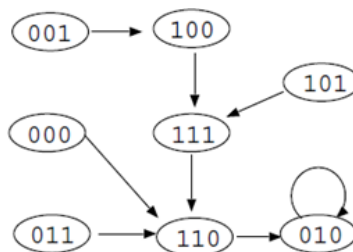
化学研究所 数理生物情報研究領域 田村武幸

背景と目的

生命活動は膨大な数の遺伝子が、細胞内で互いに複雑に制御しあうことによって維持されている。ヒトの場合には2万数千個の遺伝子が存在するが、ある遺伝子 A は別の遺伝子 B の働きを促進したり、逆に A が C の働きを抑制したりすることが知られている。DNA に記述された遺伝子の持つ情報は、RNA によって読み取られた後(転写)、翻訳されてタンパク質が生成されて機能する。ある遺伝子がどの程度働いているかは、対応するタンパク質の量(発現量)により計測される。近年、DNA マイクロアレイの技術により、各遺伝子の発現量が網羅的に測定できるようになり、遺伝子間の制御関係を数理的に表現したネットワークが得られるようになってきた。それに伴い、遺伝子制御ネットワークを数理的に解析する手法の開発が要請されている。遺伝子制御ネットワークの数理モデルとしては、ブーリアンネットワーク (BN)、ベイジアンネットワーク、微分方程式系が多く研究されている。各モデルはそれぞれ一長一短あるものの、BN は離散的な決定性の数理モデルであり、遺伝子制御の因果関係を簡潔な知識として表現するためには最適である。



ブーリアンネットワーク (BN)



状態遷移図

BN において各ノードは遺伝子に対応し、**活性化していれば1、活性化していなければ0**が割り当てられる。時刻 $t+1$ におけるノード v の値は、時刻 t における v の親ノードの値から **ブール関数で表現された制御関数**を用いて計算される。上左図は BN の例であり、 \wedge は AND、 \vee は OR、 \bullet は否定を表す。例えば時刻 t において、 $(v_1, v_2, v_3) = (0, 0, 1)$ が割り当てられたとすると、時刻 $t+1$ における各ノードの値は、 $(v_1, v_2, v_3) = (1, 0, 0)$ となる。BN の各ノードの値の

遷移は、上右図のような**状態遷移図**で表現される。よって、以後 $(1,0,0) \rightarrow (1,1,1) \rightarrow (1,1,0) \rightarrow (0,1,0)$ と遷移し、いったん $(0,1,0)$ となった後は $(0,1,0)$ から動かなくなる。この $(0,1,0)$ のように、状態遷移図において自己ループが形成される状態を(周期1の)**定常状態**という。また状態遷移図の中に長さ2以上の閉路が存在するとき、**周期的定常状態**が存在するという。

検討内容

遺伝子制御ネットワークは細胞内の遺伝子の制御関係を表現し、**代謝ネットワーク**は細胞内の化合物と化学反応の関係を表現する。**ブーリアンモデル**の遺伝子制御ネットワークの**定常状態を求める問題**、**制御する問題**や**補完する問題**は NP 困難であり、本研究ではこれらに対する**高速アルゴリズム**を開発するとともに、様々な派生する問題の**計算量解析**とアルゴリズム開発を行い、スパコンシステムを用いて検証する。

また、**GPR ネットワーク**と呼ばれる遺伝子とタンパク質、反応間の制御関係を記述する数理モデルを用いて、**遺伝子制御ネットワークと代謝ネットワークを統合**し、遺伝子の破壊や導入、細胞の初期化や状態の変化、代謝に基づく生存率や成長率を互いに関連させ**数理的にモデル化**して解析する。

また、遺伝子、反応、化合物に対し 0/1 割当の集合が与えられた時に、最小の修正を加えて矛盾を解消する問題を数理モデル化してアルゴリズムを開発し、実際の動態データを再現できるようにする (**統合ネットワークの補完問題**)。

システム生物学の分野において、微生物の遺伝子を 1～2 個ずつ網羅的に破壊し、**細胞の生存率**や**成長率**を計測する研究がある。本研究では、細胞の生存率が**化合物や反応への 0/1 割当を入力とする関数**で計算されると仮定し、このデータを用いた統合ネットワークの補完問題を解析する。

結果

バクテリアの中心代謝経路に関する統合ネットワークを、細胞生存率のデータを基に補完する問題を整数計画法で定式化し、本コンピュータシステムを用いて計算した。結果は高速かつ高精度であり、本研究結果を 2012 年 1 2 月に台湾台南市で開催された GIW2012 で発表した。