

科学研究費助成事業 研究成果報告書

平成 28 年 9 月 20 日現在

機関番号：12601

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25400195

研究課題名(和文) 生化学ネットワーク記述計算言語の意味論-局所性と因果律の解明-

研究課題名(英文) "Semantics of Computational Languages for Biochemical Networks -Elucidating Locality and Causality-"

研究代表者

浜野 正浩 (Hamano, Masahiro)

東京大学・情報理工学(系)研究科・研究員

研究者番号：50313705

交付決定額(研究期間全体)：(直接経費) 2,700,000円

研究成果の概要(和文)：RNA 転写伸張のmechano-chemicalパスウェーを確率プロセス計算の一種である rule based kappa 計算のsyntaxとsemanticsの両側面から分析した。離散確率モデルを伸張複合体の記述から導かれる化学マスター方程式によって構成し、定常状態の分析を詳細釣り合いをチェックするアルゴリズムを用いて行った。特にモデル抽象化と準定常状態近似の随伴性を示した。さらにこのパスウェーとMichaelis-Menten酵素反応との関連を介し自由エネルギーに基づく伸張の意味を与えた。

研究成果の概要(英文)：We investigated the mechano-chemical pathway of RNA Transcription Elongation in terms of computational description of stochastic process calculus. As a semantical counterpart of this, model abstraction was investigated in comparison with quasi-steady-state (QSS) approximation for chemical master equations. Thermodynamical meaning was also given on the stochastic models in terms of free energy distribution for the variously configured agents and of the traditional Michaelis-Menten enzyme kinetics.

研究分野：情報論理学

キーワード：確率プロセス計算 rule based modelling chemical master equation stochastic modelling steady state dynamics Brownian ratchet Michaelis-Menten Boltzmann distribution

1. 研究開始当初の背景

生物・化学反応のなすネットワークがなすダイナミカルなシステムの数学的記述には、伝統的に微分方程式系が使われることが多かった。この決定論的記述は、観察者を介した phenomenological な見立てによって外側からシステムを記述しようとするものであり、離散的で少数の分子の振る舞いが捨象されてしまうという問題点が残される。

この問題点は、おもいがけず情報科学で用いられていた並列計算の数学モデルである Milner の計算(90's)を用いることにより、ネットワークの構成単位である細胞分子の並列的・非同期的な相互作用に基づく内側からのきめの細かな記述により克服できるのではないかと今世紀の始めに明らかになってきた。

報告者は、これらの流れを受け、具体的な生化学ネットワークである RNA 干渉に対して、プロセス間の通信経路であるチャンネルに確率的なパラメータを付与し、確率的な相互作用を情報論理学での計算とみなす、確率プロセス計算に基づくモデリングを行い、干渉の相互作用に特徴的な非決定論的で離散的な側面を捉えることができることを明らかにしてきた。

このような新展開をみせている情報数理学と生物学のインターアクションを背景に次のような準備結果を得て当研究は計画された。

具体的な RNA 干渉ネットワークに対する Markov 過程の代数的指標と計算論的性質を対応させる一般化・拡張が、より抽象的な一次化学反応系に対して応用できることが分かってきた。これは、数理生物学者 Hans Othmer ら(2005)が与えていた化学反応マスター方程式の平均や(共)分散が遷移行列の固有値と化学ネットワークのトポロジカルな性質のみに依存して決まるという結果がおもいもかけず計算言語の確率的意味論に適用できるのではないかと当初の報告者着想による。計算言語の統語論は離散的なネットワークのグラフ的性質を相性よく記述できるため、申請者の目指す化学マスター方程式意味論に自然に反映させ組み込むことができる。

このように得られることが期待される統語論から導かれた確率的意味論に対して、(ネットワークトポロジーや代数指標を保存する)計算論的な書き換えや rule abstraction/refinement に基づく作用が定義でき、作用に対する不変量として平均や(共)分散が与えられるであろうと研究開始当初予想されていた。

2. 研究の目的

申請研究は生化学反応ネットワーク記述のためのプロセス計算体系の統語論的性質(コ

ンパクト性・rule abstraction/refinement の双対性など)を、Markov 過程や化学マスター方程式を用いた stochastic な意味論で特徴付けようとする試みであった。

具体的には、多数のプロセスが絡み合う確率的なシステムの記述に必要な不可欠な計算言語のコンパクト性や機敏性(agility)が妥当であることを、意味論上のある計算論的な作用に関する不変性(invariance)により捉えることを目的とした。

さらに生命・生物現象に典型的な多数の因子が絡み合っているながらも保たれる伝達系の秩序に構造的な説明をつけるべく、意味論を用い stochastic なモデルに特有の因果律(causality)や feed back を数学的に抽出することを目的とした。

3. 研究の方法

生化学相互作用ネットワーク記述言語に対する Markov 過程意味論を、化学マスター方程式によって構成し、時間発展に関する代数的不変量を Perron-Frobenius 指標、平均・(共)分散などにより与える。さらにこの意味論と、化学反応ネットワーク理論を関連付けることによりトポロジカルな不変量も与える。これらの不変量が言語の局所性・compact 性が妥当であることを保証し、また統語論が規定する rule abstraction/refinement の対をモデルの中での双対性をもって特徴付けることができるであろう。一方この意味論を局所的なイベントに基づく方法(グラフ書き換えや petri net)の因果律の分析に用いながら、モノダル圏の枠組みで feed back を明らかにする。さらに確率モデルに固有のゆらぎを、マスター方程式の近似としての化学 Langevin 意味論によって、情報論理的に捉える。

4. 研究成果

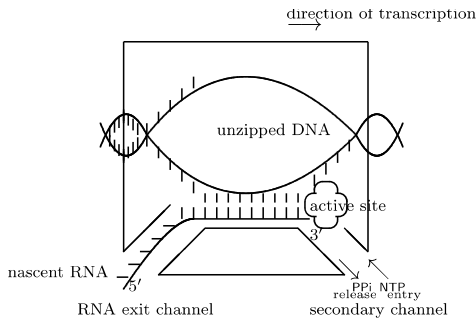
当研究では、ストカスティックな Rule based モデリングが、いかにマルチステップからなる分子相互作用をとらえられるかを解明した。このような相互作用では--確率的なゆらぎ--非 Poisson 的な待ち時間--特定の生物学的イベントの遅れ--が典型的になる。

このために、報告者は、典型的な多ステップ分子生物学的相互作用である、RNA 伸長(transcription)のモデル化を行なった。遺伝子発現の最初のステージとしての伸長は細胞内のわずか少数の分子作用本質となっており、確率的離散的モデリングが、質量作用の法則に基づく連続的決定論的な方法より適している。

報告者の方法はヌクレオチド based な内側からの記述であり、原始的なエージェントとして、1 遺伝子のなかにある数百(もしくは数千)の塩基対と(同数の)伸長中に RNA polymerase によって加えられる NTP をとった。伸長は、対応する unzipped な一本鎖 DNA をテンプレートとする転写(mRNA 合成)の主要

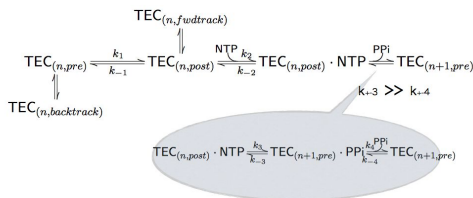
なステップである。RNAP は mechano-chemical なカップリングメカニズムであり、NTP hydrolysis からの化学的エネルギーを (back-tracked と forward-tracked 相からなる) Brownian 運動と共に機械的な仕事に変換する。この RNAP の熱力学と mechano-chemical な運動は、Transcription Elongation Complex 伸長複合体 (TEC) を媒体とする。TEC は RNAP とテンプレート DNA と発生する RNA の組み合わせで形つくられる。

図 1 TEC



TEC の主要な熱力学的機能はヌクレオチドの合併と TEC の DNA 上の 1bp 前進により誘導される、NTP binding に関する mechano-chemical な循環性である。これは以下の生化学的な Brownian ラチェットパスウェイで知られており ((n,*) の n は転写サイズで *-translocation 相)、化学エネルギーを mecanochemical エネルギーに変換するかを示している。

図 2 ラチェットパスウェイ

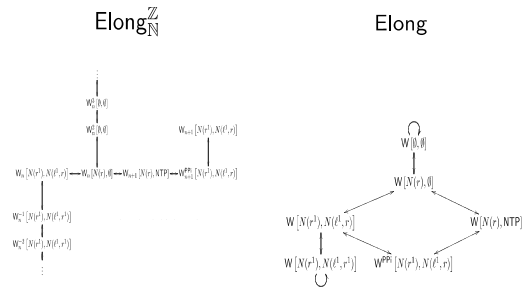


(1) TEC のコンパクトな活性部位による Rule Based な統語論的記述:

ヌクレオチドを単量体であり 2 つのリン酸化ポンドを持つ部分ユニットとして、また TEC を複数個部分の活性化したサイトからなる窓枠を表すエージェントとして統語論的に表現した。窓枠エージェントは RNA と DNA に沿って滑走する。これらのエージェントの組み合わせにより TEC の post-translocation と pre-translocation の 2 つの相を区別できる。

このエージェントの定式化によりラチェットパスウェイに対応する規則たちが、きめの細かいものから粗いものまで、エージェントの特定のサイトを忘却することにより、一様に得られる。

図 3 ラチェット規則と抽象化



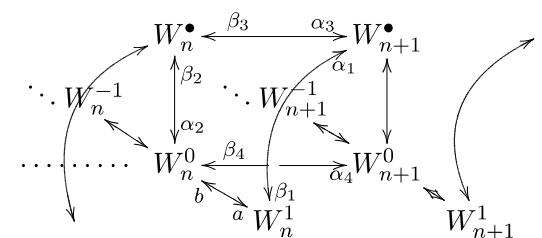
逆に、特定のサイトをエージェントに加えることにより、(粗いものから)微細な規則も得ることができた。これらサイトを忘却することにより得られるモデル抽象化の健全性は後記 2 の手法で明らかになる (図 5 参照)。

(2) マスター方程式意味論と抽象化:

化学マスター方程式を用いて上記 1 の確率離散意味論を構成した。この意味論は、化学反応から得られる Kolmogorov 方程式の一種であり、Markov 過程の時間発展を規定する。この方程式の熱力学的極限として、質量法則に基づく連続的な速度方程式が得られる。情報論理学の観点からの報告者の手法は少数のエージェントのふるまいやゆらぎを捨象せずに記述することを指向するもので、当研究にとってこの確率的離散意味論が本質的となる。一般的にはマスター方程式系を解析的に解くことは困難であるが、報告者のケースでは、TEC の相互作用が局所的であり、分子の多数性によらないことより、方程式系は有限次元とまることが分かった。この局所性は、生物学的には、TEC の DNA テンプレートと RNA に関する熱力学的安定性から生じている。

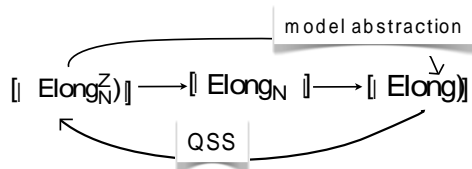
TEC の窓枠からのみ見える局所的な相互作用に基づく 1 で得られた規則は、周辺のエージェントの大局的な状態よらないものであることから、1 の統語論記述の局所性により、(エージェントのサイトの状態に関する組合せ爆発をおこさない)コンパクトなマスター方程式系が得られた。この方程式の定常分布を化学者が詳細釣り合いをチェックする Wegscheider アルゴリズムを用い分析した。

図 4 $\llbracket \text{Elong}_N^Z \rrbracket$ が規定する状態遷移



さらに、マスター方程式意味論の間の抽象化の健全性が、生化学で用いられている Quasi-Steady-State-approximation (準定常状態近似)として実現できることを示した。これは化学反応速度に関するある仮定のもと、きめの細かさをどう回復できるかを保証できるものであり、情報論理的には、モデル抽象化との随意性により特徴づけられた。

図 5 抽象化と QSS の随伴性



(3) 転写伸長の熱力学的意味 :

上記 1 と 2 でそれぞれ構成された統語論と意味論の対応の中で各エージェントの組み合わせ対する自由エネルギープロファイルを、Boltzmann 分布(エネルギーと確率の対応)を介して与え、化学平衡状態の分析を行った。これは TEC の Brownian ラチェットパスウェイの主要部分を Michaelis-Menten 型の酵素反応と見なせるとする報告者の着想から可能になり、これから特に化学平衡の特徴付けが計算モデルの枠組みでできた。

図 6 Boltzmann と Michaelis-Menten

$$v = \frac{v_{max}[NTP]}{K_n + [NTP]} \quad \text{with} \quad K_n = \frac{k_{-2}}{k_2} \log \prod_j \frac{w_n^{[j]}}{w_n^{[0]}}$$

$$\sum_j \exp((\mathcal{E}(W_n^0) - \mathcal{E}(W_n^j))/k_B T)$$

energy difference

v は TEC(n+1,pre) の速度

(4) 今後の展開

RNAP は分子モーターであることを考慮すれば、報告者によって得られた Rule based な Brownian ラチェットの記述にどのように drift (熱的力) を組み入れるのかが今後の課題として残された。報告研究の確率フラックスは拡散に関する Fick の式に drift 項を重ね合わせ得られることから、言い換えれば、どのように Fokker-Planck 式が計算言語を使ってモデルかできるのか? が今後の課題である。

5 . 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 1 件)

Masahiro Hamano, "Stochastic Transcription Elongation via Rule Based Modelling", 査読有, Electronic Notes in Theoretical Computer Science 18639, Elsevier, DOI: 10.1016/j.entcs.2016.09.019

[学会発表](計 1 件)

Masahiro Hamano, "Stochastic Transcription Elongation via Rule Based Modelling", Sixth International Workshop on Static Analysis and Systems Biology (SASB 2015), Saint-Malo, France, 2015 年 9 月.

6 . 研究組織

(1) 研究代表者

浜野 正浩 (Masahiro HAMANO)

東京大学大学院、情報理工学系研究科、
研究員

研究者番号 : 50313705