

時間論理による仕様記述を用いた遺伝子ネットワークの定性的振る舞い解析

伊藤 宗平†

泉 直子‡

萩原 茂樹†

米崎 直樹†

†東京工業大学大学院情報理工学研究科計算工学専攻

‡十文字学園女子大学社会情報学部社会情報学科

1 はじめに

本研究では、遺伝子ネットワークの可能な振る舞いを定性的手法を用いて総括的に捉え、その中に特定の性質を満たすものが存在するか、あるいは全ての振る舞いは特定の性質を満たすかということを解析する方法を与える。可能な振る舞いの制約を、システムの動作が満たすべき“仕様”とみなし、時間論理による仕様記述からその充足可能性を判定することでモデルを構築することにより、仕様検証の枠組みを用いて上記の解析を行う。

2 関連研究

生物システムの解析を時間論理を用いて行う研究は近年よく見られるようになってきている。それらは主にモデル検査手法 [1] を用いるものであり、定量的なパラメータを与えられた生物系のモデル（微分方程式や化学反応式など）から、それをオートマトンやマルコフ連鎖などで抽象化し、その上で調べたい性質をモデル検査するというものである ([2, 3] など)。ただし、この手法は生体内における反応速度などのパラメータが分からなければモデルが与えられないため適用することができない。そのため、パラメータを使わずに、生物系を定性的に解析する手法も必要である。定性的な手法では、パラメータや初期設定や外部入力のあらゆる可能性を考慮した可能な振る舞い全てを取り扱い、その中に特定の性質を満たすものがあるか、またどのような振る舞いでも必ず満たされる性質は何か、といった質問に答えることができる。こういった研究はいくつかあるが ([4, 5] など)、どちらも可能な振る舞いの記述力が弱く、またシステムへの外部入力の影響を無視していた。本研究では、可能な振る舞いの集合を規定する制約を記述する言語として線形時相論理 LTL [6] を採用し遺伝子の発現状態や産物の濃度レベルに関する適切な命題を導入することで、遺伝子ネットワークの振る舞いの柔軟な記述法を得た。また、本手法では外部入力の影響を考慮した記述も可能である。

3 本研究の手法

遺伝子は発現することによりたんぱく質を産生するが、その産物が他の遺伝子のプロモータ領域などに結合することにより、他の遺伝子の発現を活性化あるいは抑制することがある。その遺伝子間の活性・抑制の関係をネットワークとして表したものが遺伝子ネットワークである。遺伝子調節では、調節遺伝子の産物がある一定の濃度以下であるときには調節効果はほとんど現れず、その濃度を超えると急激に効果を表すという現象が知られている。これは図 1 のように、S 字の曲線として表現される。

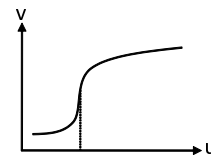


図 1: u が v を活性化するときの濃度の関係。

このことから、 u の発現量がある一定値（閾値）を超えた時に v は ON になり、そうでないときは OFF であると捉える事が出来る。状態が ON のときの遺伝子は発現しており、その産物の濃度は上昇していく。OFF のときは発現が止まっており、産物は時間とともに分解していくので濃度は減少していく。これが遺伝子ネットワークの振る舞いの基本原理であり、このことを適切な命題を用いて LTL で表現することにより、可能な振る舞いの満たすべき制約が得られる。本手法では、各遺伝子ごとにそれが ON であるかどうかを示す命題と、各遺伝子の発現量の濃度レベルを表す命題を導入した。ネットワークのパターンをどのような制約として記述すべきかについては、紙面の都合上掲載できないので著者らの論文 [7] を参照されたい。

本研究の提案する解析手法では、まず対象となるネットワークの可能な振る舞いの制約式 φ を記述する。次に、対象となるネットワークに対して調べたい性質 ψ を LTL で記述する。この時、可能な振る舞いの中にその性質を満たすものが存在するということは、 $\varphi \wedge \psi$ が充足可能であることに対応し、可能な振る舞い全てがその性質を満たすということは、 $\varphi \rightarrow \psi$ が恒真であることに対応する。どちらも LTL の充足可能性判定問題に帰着

Qualitative analysis of gene network by temporal logic specification of behaviors

†Sohei ITO ‡Naoko IZUMI †Shigeki HAGIHARA †Naoki YONEZAKI

†Department of Computer Science, Graduate School of Information Science and Technology, Tokyo Institute of Technology

‡The Department of Social and Information Sciences

する。

次の節では、緑膿菌の粘液生成の問題に本研究の手法を適用した例を示す。

4 適用例

緑膿菌は自然環境中に存在する日和見感染細菌である。この細菌は嚢胞性線維症の患者の肺においてのみ粘液(アルギン酸)を生成するが、これは緑膿菌の抗生物質耐性を高め、患者の呼吸不全を引き起こす。アルギン酸生成に関する遺伝子ネットワークは図2のようになっており、 z はアルギン酸の生成を表すノードである。

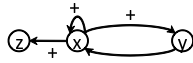


図2: 緑膿菌のアルギン酸生成に関する遺伝子ネットワーク。 $x=AlgU$, $y=anti-AlgU$, $z=alginate\ synthesis$ 。

緑膿菌は、自然環境下では粘液を生成しないが、一旦粘液生成状態になったなら、患者の肺から取り出してもその状態を維持するという特性が知られている。このような性質をこの遺伝子ネットワークがもつか否かを本手法を用いて調べる。

留意すべき命題は、 $\{on_x, on_y, on_z, x_x, x_y, x_z, y_x\}$ である。 on_x, on_y, on_z は各遺伝子が ON かどうかを表し、 x_x, x_y, x_z はそれぞれ x の濃度レベルが x, y, z を活性化するための閾値を超えているかどうかを表す。同様に y_x は y の濃度レベルが x を抑制するための閾値を超えているかどうかを表す。これらの記号を使ってこのネットワークの可能な振る舞いの制約を記述する。その一部を以下に示す。

$$G(x_z \leftrightarrow on_z) \quad (1)$$

$$G((x_x \wedge \neg y_z) \rightarrow on_x) \quad (2)$$

$$G((\neg x_x \wedge y_z) \rightarrow \neg on_x) \quad (3)$$

$$G((on_x \wedge x_x) \rightarrow (x_x W \neg on_x)) \quad (4)$$

$$G((\neg on_x \wedge \neg x_x) \rightarrow (\neg x_x W on_x)) \quad (5)$$

$$G(x_z \rightarrow (x_x \wedge x_y)) \wedge G(x_y \rightarrow x_x) \quad (6)$$

(1)-(3) は遺伝子が ON, あるいは OFF になる条件である。各遺伝子の調節関係から導出される。(4) は遺伝子が ON のときに濃度レベルの上昇を表した式である。(5) は遺伝子が OFF のときに濃度レベルの下降を表した式である。各濃度レベルごとに(4)(5)のような式が与えられる。(6) は濃度レベルの全順序を表している。

次に、このネットワークについて調べたい性質を LTL

で記述する。

$$on_z \rightarrow GFon_z \quad (7)$$

$$\neg on_z \rightarrow G\neg on_z \quad (8)$$

(7) は一旦粘液生成状態になったなら、定常的に粘液を生成するという性質で、(8) は自然環境下では粘液を生成しないという性質である。

これらすべての式の連言をとり、その充足可能性を判定した結果、充足可能であるという結果を得た。このことにより、このネットワークは既知の事実に適合したものであることが分かる。

5 まとめ

本研究では定性的に遺伝子ネットワークの振る舞いの解析を、LTL による仕様記述から充足可能性を判定することで行う手法を提案した。今後の研究方向として、部分構造のモジュール化を利用した部分検証の手法を遺伝子ネットワークの解析に導入することや、性質を満たすためにネットワークにどのような修正を行えばよいかを示唆する手法の開発などが考えられる。

参考文献

- [1] E. Clarke, O. Grumberg, and D. Peled. *Model Checking*. MIT Press, 1999.
- [2] G. Batt, D. Ropers, H. de Jong, J. Geiselman, R. Mateescu, M. Page, and D. Schneider. Validation of qualitative models of genetic regulatory networks by model checking: Analysis of the nutritional stress response in *Escherichia coli*. *Bioinformatics*, Vol. 21, No. Suppl.1, pp. i19–i28, 2005.
- [3] M. Kwiatkowska, G. Norman, and D. Parker. Using probabilistic model checking in systems biology. *ACM SIGMETRICS Performance Evaluation Review*, Vol. 35, No. 4, pp. 14–21, 2008.
- [4] Franois Fages, Sylvain Soliman, and Nathalie Chabrier-rivier. Modelling and querying interaction networks in the biochemical abstract machine biocham. *Journal of Biological Physics and Chemistry*, Vol. 4, pp. 64–73, 2004.
- [5] G. Bernot, J.P. Comet, A. Richard, and J. Guespin. Application of formal methods to biological regulatory networks: extending thomas' asynchronous logical approach with temporal logic. *Journal of Theoretical Biology*, Vol. 229, No. 3, pp. 339–347, 2004.
- [6] E. Allen Emerson. Temporal and modal logic. In *Handbook of Theoretical Computer Science, Volume B: Formal Models and Semantics (B)*, pp. 995–1072. 1990.
- [7] Sohei Ito, Naoko Izumi, Shigeki Hagihara, and Naoki Yonezaki. Qualitative analysis of gene regulatory networks by satisfiability checking of linear temporal logic. In *Proceedings of the 10th IEEE International Conference on Bioinformatics & Bioengineering*, pp. 232–237, 2010.