

文献・公開データからの知識抽出による スマートセル設計支援技術

(1. 株式会社日立製作所、2. 国立研究開発法人 医薬基盤・健康・栄養研究所、3. 京都大学大学院医学研究科) ○伊藤 潔人¹・荒木 通啓^{2,3}

Smart cell design support technology by knowledge extraction from literature and public data
(¹Hitachi, Ltd., ²National Institutes of Biomedical Innovation, Health and Nutrition,
³Graduate School of Medicine, Kyoto University) ○Kiyoto Ito,¹ Michihiro Araki^{2,3}

In the design-build-test-learn (DBTL) cycle of microbial developments, many of the interpretation of various data and the design improvement based on it depend on a researchers' individual knowledge and the manual investigation of literature and public database. Such knowledge acquisition process limited the speed of the microorganism development and was a big problem which should be solved technically such as the systematic knowledge accumulation, the discovery, and the reuse were difficult. On the other hand, recent advances in machine learning and artificial intelligence (AI) technology have made it possible to extract new knowledge and design patterns from literature and public database. Therefore, by applying such technology to DBTL cycle, to solve the above problems, it is expected to further accelerate microbial development. In this lecture, we will introduce a machine learning technology that focuses on knowledge extraction techniques and enzyme gene search from literature.

Keywords : Metabolic design, Microbe, Text mining, Machine learning

微生物開発の DBTL サイクル (Design-Build-Test-Learn) において、各種データの解釈と、それをもとにした設計改良の多くは、個人の知識背景や、手作業による文献・データベースの調査に依存している。こうした属人的な知識獲得プロセスは、微生物開発において律速であるとともに、体系的な知識蓄積・発見・再利用が困難であるなど、技術的に解決すべき大きな課題となっていた。一方、近年の機械学習・人工知能 (AI) 技術の進展に伴い、文献や公開データから新たな知識・パターン抽出も可能となってきた。そこで、こうした技術を DBTL サイクルに応用することで、上述の課題を解決し、微生物開発を更に加速することが期待されている。本講演では、こうした背景を踏まえて開発された、文献からの知識抽出技術と酵素遺伝子探索にフォーカスした機械学習技術を紹介する。

