

⑱ A Machine Discovery from Amino Acid Sequences
by Decision Trees over Regular Patterns

S.Arikawa, S.Kuhara, S.Miyano*, Y.Mukouchi, A.Shinohara, T.Shinohara(九州大学, 日本)

発表要旨

この講演では、アミノ酸の列からtransmembrane domainを同定する問題に対する解を、効率良く発見することができた機械学習システムについて述べられた。

そこで用いた方法は、Quinlanのdecision treeに関するID3の方法を、形式言語の分野では馴染みの深いregular patternを用いて実現したものであり、正規言語族の真部分集合となる言語族の新しい受理方法が提案された。

この言語族の学習可能性について、変数の個数を高々 k に制限した k -variable regular patternを用い、treeの深さを高々 d に制限した場合には、PAC学習の枠組で多項式時間計算可能であるという計算量的基礎付けも述べられた。

また、具体的なアルゴリズムも示され、そのアルゴリズムを用いて行なわれた実験では、この問題に対する非常に良い精度の、生物学的にも興味深い近似解を効率良く求められたことが報告された。

質疑応答

質問：PAC学習についての結果を示しているが、帰納推論の意味で多項式時間学習可能性との関係ははっきりしていないのか？ 多項式時間PAC学習可能なクラスは一般に帰納推論での多項式時間学習可能なクラスより小さいように思えるが。

回答：はっきりしていないが、関係を明らかにすれば、応用の可能性が広がると考えられる。

質問：その仮説についての生物学者のviewはどうか？

回答：生物学者のviewと一致していた。

質問：最後の仮説は新しいものなのか？

回答：新しいものである。