

論文題目 形式文法の確率的一般性と応用に関する研究

氏名 柴田 剛志

正例から形式文法を帰納推論する枠組みが導入されて以来、様々なクラスの学習可能性が考えられてきた。正例からの学習可能であるために文法が満たすべき性質の必要十分条件も与えられている。正例からの学習はもっとも基本的な帰納推論の枠組みであるが、負例や質問などを用いるもの比べて、学習可能なクラスは少ない。例えば正規言語のクラスなど正例からの学習できないことがわかっている。また、たとえ学習可能であることが示せても、効率的な学習アルゴリズムが存在するかどうかは別の問題であり、一般に、言語クラスの豊かさと、学習の効率性の間にはトレードオフが存在すると考えられる。

近年、正例からの学習の分野において、文脈自由文法の非正則なサブクラスで、効率的に学習可能なクラスが発見されてきた。Very Simple Grammarのクラスや、その拡張である Right-unique Simple Grammar のクラスがそうである。両者はともに Simple Grammar とよばれる文脈自由文法のサブクラスのサブクラスである。これらの効率的学習アルゴリズムに対していろいろな応用を考えることができる。

そのひとつとして、本論文では強化学習へ応用する方法を提案している。強化学習の理論的な基礎となっている確率モデルはマルコフ決定過程と呼ばれるものである。さらに、観測状態と真の状態が一致しないような場合は部分観測マルコフ決定過程と呼ばれており、マルコフ決定過程のクラスはそのサブクラスである。一方で、有限状態マルコフ決定過程にスタックを付与した決定過程を考えれば、観測の履歴が真の状態にそのまま対応しないため、履歴からでは、どのような状態とスタックの構造をしているのかは直接推測することはできない。しかし、このような状態とスタックの組は、Right-unique Simple Grammar とみなすことができ、観測の履歴は、Right-unique Simple Language の正例となることがわかる。したがって、Right-unique Simple Grammar の学習アルゴリズムを用いれば、観測の履歴、即ち正例から帰納的にもとの文法、即ち状態とスタックの組を推測することができる。しかし、この場合、言語が等しいだけでは正しい文法を帰納推論としたとすることができず、確率言語の意味で齟齬をきたしえないような文法を出力しなければならない。ここで注意すべきは、毎回の履歴が観測される確率は行動によって左右されるため、行動を抜き考えると、毎回の履歴は独立

ではなくなり、行動が多数存在したり連続的だったりした場合、PAC学習などの標準的な枠組みで扱うことは難しいことである。しかし、次に述べる結果を用いれば、そのような確率的な齟齬を除くことが可能である。

本論文では、形式文法、とくに Simple Grammar に対する確率的一般性を、その文法の生成規則に確率値を割り当てた場合に生成しうる確率言語全ての集合ととらえ、それらの間の包含関係について議論を展開している。より具体的に書くと、文法  $A$  にたいし、それが生成しうる確率言語全ての集合を  $K(A)$  とする。ある文法クラス  $X$  に属する任意の文法  $A, B$  に対して、 $K(A) \subset K(C)$  かつ  $K(B) \subset K(C)$  となる文法  $C$  がクラス  $Y$  に存在するとき、クラス  $X$  はクラス  $Y$  の範囲内で統合可能であると呼ぶことにする。本論文では、Very Simple Grammar のクラスがそれ自身の範囲内で統合可能であるのにたいして、Right-unique Simple Grammar (以下 RSG) や Simple Grammar (以下 SG) のクラスはそれ自身の範囲内では統合不可能であることを示している。また、新たに、RSG のクラスがそのサブクラスとなる、Unifiable Simple Grammar という文法を定義し、その文法クラスがそれ自身の範囲内で統合可能であることを示すことで、RSG のクラスが SG のクラスの範囲内で統合可能であることが示される。ことことから、RSG の特性を考慮すれば、RSG のクラスが「強い意味」で SG のクラスの範囲内で統合可能であることが示せる。ここでいう「強い意味」とは、任意の RSG  $A$  に対して、ある SG  $B$  が存在して、 $A$  と言語が等しい任意の RSG  $C$  に対して  $K(C) \subset K(B)$  となることである。また、実際にそのような  $B$  を構成する方法、及び構成するための計算量、及び構成後の文法のサイズを与えている。