

研究タイトル：

# パターン情報処理と学習システム



氏名：	平岡和幸／HIRAOKA Kazuyuki	E-mail：	hiraoka@wakayama-nct.ac.jp
職名：	准教授	学位：	博士(工学)
所属学会・協会：	日本神経回路学会, 日本応用数理学会		
キーワード：	パターン認識, 機械学習		
提供可能技術：	・パターン情報処理		

## 研究内容：

複雑で解析が難しい課題、前もって完全には特定できない部分を含む課題、時間がたつにつれ変化する課題などでは、システムの適切な出力を使用者が指定することは困難もしくは煩雑である。そのような状況でも運用が可能なシステムを実現するには、試行錯誤的な自動学習を行う能力をシステムに付与することが求められる。強化学習は、この型の学習問題を扱う枠組である。

Q 学習をはじめとする従来の強化学習法の大部分は、単一のタスクを想定している。これに対し本研究では、複数のタスクからなるタスク族を想定し、それぞれに対する最適な行動則を一括して学習する手法を提案する。

一般に、強化学習タスクは遷移則と報酬則により記述される。本研究で扱うタスク族は、共通の遷移則を持ち、報酬則の異なるタスクからなる集合である。具体的には、報酬が複数の部分報酬の荷重和となる荷重報酬モデルを仮定し、荷重を変えることで得られる無限通りのタスクすべてを学習の対象とする。

本研究で提案する手法は、荷重報酬モデルの全タスクを有限の計算量で一括学習することができる。しかも、その結果は、無限通りの各タスクに対してそれぞれ個別に Q 学習を行った結果と等しい。このようなことが可能となるのは、モデルが持つ二つの著しい性質を活用するからである。それらの性質を根拠に、行動価値関数を荷重の区分一次関数に限定し、全荷重に対する行動価値関数を有限個のパラメータで表現することが、提案手法の鍵である。この表現のもとでは、Q 学習の更新処理は凸包の併合と Minkowski 和に帰着される。凸包や Minkowski 和の計算法は計算幾何学の分野で深く研究されており、よく知られた効率的なアルゴリズムを利用することができる。

タスク族に対する一括強化学習には、次のような応用が考えられる。

- ・時変環境：時間がたつにつれ断続的にタスクが変化する場合、現在経験しているタスクの学習と並行して、将来出現し得るタスク族に対しても裏学習を進めておけば、タスクが変化しても即座に対応が可能となる。
- ・多目的問題：両立しない複数の指標の向上を求められる場合、簡便な対処として、全指標の荷重和をとることで単一の指標に帰着させる方法がしばしば用いられる。しかし、指標間の選好度合を荷重という形式でしかも事前に表明することは、実際には容易でない。提案手法では、あらゆる荷重に対する学習が一括して進められる。このため、学習中や学習後でも、結果を見比べながら荷重を自在に選択し直すことができる。

## 提供可能な設備・機器：

名称・型番(メーカー)	