

# 論文内容の要旨

論文題目 “State-Action Map Compression by using Vector Quantization for Decision Making of Autonomous Robots”

(自律ロボットの行動決定のための状態行動地図のベクトル量子化圧縮)

氏名 上田 隆一

本論文では、ロボットの行動決定則（方策）をロボットのランダムアクセスメモリ（RAM）に小容量で記述する手法を提案する。提案する手法では、動的計画法（dynamic programming, DP）で計算され、配列上に記述された行動決定則（状態行動地図）を、ベクトル量子化（vector quantization, VQ）で圧縮するという手順により、小さいメモリ消費で効率の良い方策を作成する。

1章では、研究背景、関連研究、本論文の目的について述べる。本論文ではロボットの行動決定問題としてマルコフ決定過程と最適制御問題を扱うため、それらを定式化して方策の最適性に関して一般的な説明を行う。その後、ある問題に対して次のような条件を満たす方策を作成するという問題を提起する。条件は、1) 方策中のロボットの行動が、データ読み出し手続きのための演算のみで読み出せること（反射的方策であること）、2) 最適制御問題の観点から効率が良いこと、3) 記述に必要なメモリ量が少ないこと、の3点である。このような方策は、計算資源に乏しいロボットや、実時間で行動する必要があるロボットにとって非常に有用である。このような問題に対する関連研究は非常に多数存在するが、実装時のメモリ容量に注目した研究は行われていない。関連研究の多くでは、行動決定問題を解く際に必要な関数近似手法に注目しており、関数を表現するための要素数を減らすことに焦点が当てられている。これらの研究で提案された手法は方策のメモリ量削減にも利用できる一方、要素あたりのメモリ使用量が多くてむしろ方策のメモリ量を増大させるなどの副作用も存在する。換言すれば、既存のどのような手法において得られた方策も、更に小さくできる余地があるということである。このような状況を鑑み、本論文ではメモリ量（ビット）に対して効率が良い反射型方策を作成する方法論を提案することを目的とする。

2章では、反射型方策の表現方法として最良な、ルックアップテーブル上に記述された方策（状態行動地図と呼ぶ）について定義し、作成方法について説明する。状態行動地図は、ロボットや環境のあらゆる状態に対して分類、番号付けし、その番号順に適したロボットの行動を記述したバイナリ列として定義される。さらに、状態がいくつかの変数（状態変数）を元とする変数で表現されるとき、状態行動地図は状態空間中の多次元配列として定義される。状態行動地図の作成方法として、本論文ではDPの

実装方法の一つである価値反復アルゴリズムを用いるため、このアルゴリズムについて説明する。

状態行動地図のような単純な配列表現はメモリ量の点で非効率で無策な方法として見られがちである。このような批判はある程度正しい。しかし本章では、ある例においては他の有名な方策表現方法よりも、状態行動地図の方がビット当たりの効率が良いことを明らかにする。

本章から4章までは、手法の説明と評価のために水たまり問題を使用する。水たまり問題は、人工知能の評価のための標準問題として良く知られている例題である。

3章では、ベクトル量子化(VQ)による状態行動地図の圧縮手法が提案される。本手法を適用することで状態行動地図のメモリ量を小さくすることができる。ただしこの際、不可逆圧縮を用いるため、圧縮された地図(VQ地図と呼ぶ)の効率は下がる。VQは主にデジタル画像や音声の圧縮に用いられる。この際、不可逆圧縮で生じるデータの差異(劣化)を数値化する尺度:歪み測度が定義され、歪み測度を最小化するように圧縮が試みられる。デジタル画像や音声の場合、歪み測度は各要素の数値の変化量で簡単に定義することができる。一方、状態行動地図の場合は一箇所の行動の変化が状態行動地図中のあらゆるところに連鎖的に影響を及ぼすため、歪み測度の定義、あるいは圧縮自体が非常に困難となる。そこでVQを状態行動地図の圧縮に適用するために、新たな歪み測度:状態価値歪みを提案し、上記で述べた困難さのなかでも効率の良いVQ地図を作成することを試みる。本章では水たまり問題において、作成されたVQ地図と他の方策表現(圧縮前の状態行動地図、木構造で表現された方策)との比較を行う。

4章では、VQ地図のビットあたりの効率を向上させるための手法を提案する。特に、VQ地図の価値反復とVQ地図のVQによる再圧縮については水たまり問題において定量的な評価を行い、3章のVQ地図と比較する。他にも、VQにおいてベクトルを定義する際に圧縮に適した方法を推定する手法、状態行動地図を数パーティションに分けて別々にVQを適用する手法について説明する。

5章では、アクロボットに対して提案手法を適用し、評価を行う。アクロボットは2リンクに対して1つのアクチュエータを持つ劣駆動マニピュレータの一種である。本章ではアクロボットの振り上がり問題(ある高さまでの最短時間到達問題)を扱う。アクロボットの制御は、水たまり問題とは異なる性質を持つ。特に、その非線形で複雑な挙動が原因で、制御問題を解くことが困難とされている。しかし本章では、提案手法を水たまり問題と同様にマルコフ決定過程として定式化し、提案手法を水たまり問題と同じ方法で適用できることを示す。評価シミュレーションでは、提案手法が4次元配列の状態行動地図を2次元のVQ地図まで次元を落として圧縮できることを示す。

6章では、提案手法をロボットサッカー(ロボカップ)に適用する。ロボカップは、現在最も有名なロボット工学・人工知能の標準問題となっている。本章では、2種類の異なる性質を持つタスクを定義し、提案手法を適用する。

一方のタスクは、ロボット(SONY製4足ロボットERS-210)がボールに最小歩数で到達するという

問題を扱い、実機実験を行う。この実験では、1.5[%] のビット数まで圧縮された VQ 地図が、実機においてもとの状態行動地図とほぼ同じ効率を有することを確認する。

もう一方のタスクでは、2 台の ERS-210 が最短時間で得点する（ボールをゴールに入れる）問題をシミュレーションにおいて扱う。マルチエージェント系は DP を含めて提案手法にとっては状態変数の多さから、最も不向きな問題である。マルコフ決定過程として定式化したところ、状態行動地図の要素数は 6 億状態に達した。このような規模の問題は DP では解くことが非常に困難であり、また状態行動地図のメモリ量は、ERS-210 の RAM (16[MB]) よりも大きくなる。本章では、そのような規模の問題を、比較的高性能なデスクトップコンピュータ (3.2GHz Pentium D と 3.0GB の RAM 搭載) で価値反復によって解き、状態行動地図を得ることに成功した。さらに、ERS-210 に搭載可能なメモリ量の VQ 地図を作成することに成功した。そして、シミュレーション上ではあるが VQ 地図がロボットを協調させて効率よくボールをゴールに運ぶことを確認した。この結果は、提案手法が計算資源の少ないロボットに効率の良い反射型方策を与えられることを示していると考えられる。

7 章では、これまでの適用例を踏まえて、提案手法中の各アルゴリズムや概念について評価、議論を行う。特に、状態価値歪みの汎用性について詳しく議論する。

8 章では、本論文の結論を述べる。さらに、提案手法に対して様々に考えられる拡張方法について説明を行う。