

# EMERGENT BEHAVIOR

## 連続的観測空間に対する部分観測環境下における自律的状态空間構成

本研究グループでは、不完全・不確実で連続的なセンサ入力に基づくロボットが、環境との相互作用に基づいて適切な状態識別機構を自律的に獲得して行動学習を行うための学習手法を提案している。

ロボットをはじめとする身体性を持ったエージェントにとって、自らのセンサを通して得られる知覚入力を、行動を出力するために十分正確な状況認識へ変換するプロセスは本来トリビアルなものではない。したがって、設計者による事前知識の埋め込みを廃した自律的な行動を実現するためには、エージェントはこの変換様式を環境との相互作用を通して獲得しなければならない。

これを実現する上で解決せねばならない本質的問題は以下の2点である：(1) その局所性とノイズによってマルコフ性を失っているセンサ入力を手がかりにして、マルコフ的状态量の推定を行う、(2) 連続的・多次元のセンサ入力を、タスクに対する意義に即して自律的に切り分ける。本研究では、短期記憶（過去の観測値・行動の系列）に基づく識別を表現する決定木構造の状態表現 (Figure 1) を用いることにより知覚入力の非マルコフ性に対応する。この状態表現を、タスク実行において実際に得られた経験に基づいて逐次的に分割することによって、タスクに対して必要十分な状態識別を実現する状態表現を目指す。また、状態表現上の観測値に対応する空間に適切な分割を加えることにより、設計者による事前のセンサ空間の設計なしに、センサの解釈を自律的に実現する。

検証シミュレーションとして、Figure 2 に示す環境におけるナビゲーションタスクを行った結果、エージェントは適切な状態表現を構成し、最適な行動が獲得された (Figure 3)。

*Keywords:* Partially Observable Markov Decision Process (POMDP), Autonomous State-Space Segmentation.

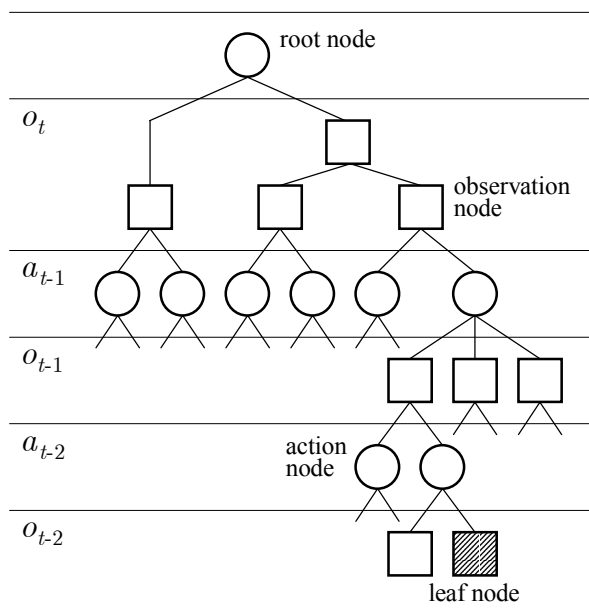


Figure 1: State-representation

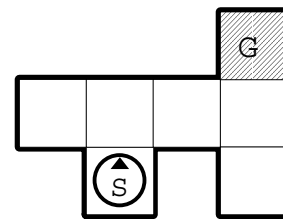


Figure 2: Simulation environment

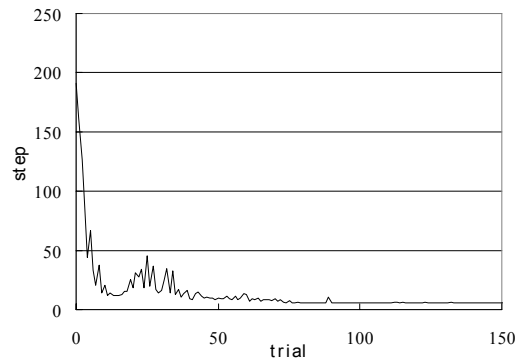


Figure 3: Performance improvement