

センサ数の過不足が学習に及ぼす影響について

室蘭工業大学 ○中南義典, 室蘭工業大学 倉重健太郎, 株式会社本田技術研究所 尾上由希子

要旨

ロボットはセンサの認識情報を用いて学習を行う。そのためロボットが持つセンサが異なると学習効率も異なると考えられる。しかし学習手法の有用性を検証する際、センサ数に過不足が生じたまま検証が行われている可能性がある。このときセンサ数の過不足により学習効率に変化し、検証が適切に行われていない可能性がある。本論文ではセンサ数の過不足が学習に及ぼす影響を調査し、学習に合わせたセンサ選択の重要性を示す。

1. はじめに

ロボットが学習を行う時、自身に搭載されたセンサから得られた情報を元に自身の状態を認識して学習を行う。そのため、タスクに対してセンサ数が不足しているロボットは適切な学習を行うことが難しい。一方、ロボットに多くのセンサを搭載してセンサから得られる情報量が増えると、より多くの情報に対して学習することになり学習にかかる時間が増えてしまう。このように、ロボットの学習においてはセンサ数の過不足が学習効率に大きく影響を与えると考えられる[1, 2]。

従来研究においては、ロボットに与えるタスクに対して必要と考えられるセンサを研究者が予測し、ロボットに搭載するセンサを選択している[3]。しかし、ロボットが実際に過不足なくセンサを搭載していることを判断する明確な指標は無い。つまり、ロボットのセンサ数に過不足が生じたまま研究が行われている可能性がある。このときもし、センサ数の過不足によって学習効率に変化が生じるならば、その影響が見過ごされてしまう。このような状況では、学習手法の有用性を検証する際にセンサ数の過不足により本来の学習効率が発揮されず、学習手法が本来持つ有用性が見過ごされてしまう可能性がある。

そこで本論文では、ロボットのセンサ数の過不足が学習に及ぼす影響について、強化学習[4]を対象として調査を行う。この調査により、ロボットの学習において学習に合わせたセンサを搭載することの重要性を示す。

2. 調査対象となるセンサの定義

本論文ではいかなるセンサにも共通して起こる学習への影響を調査する。そのために、センサの基本的な機能について抜き出して抽象化したセンサを定義する。この抽象化したセンサに対してシミュレーション実験を行うことで、センサが学習に及ぼす影響を調査する。

まず、センサの基本的な役割について考える。あらゆるセンサに共通する役割は、対象の状態を認識し、認識した状態を離散値に変換して出力することである。そこで本論文では、認識した状態を式(1)に示す離散値として出力するものをセンサとして定義する。

$$V = \{1, 2, \dots, V_{\max}\} \quad (1)$$

エージェントはこのセンサの一つ、または複数用いて状態を認識する。エージェントは、自身が持つ各センサから出力された値の組み合わせによって自身の状態を決定する。そのため、エージェントが認識する状態空間は自身が持つセンサの組み合わせによって決まる。

次に、エージェントが持つセンサの過不足について考える。エージェントが必要とするセンサは、エージェントに与えるタスクによって決まる。そこで、エージェントに与

えられたタスクの達成に必要なセンサを必要センサと呼び、タスクに対して必要センサが少ないエージェントはセンサ数が不足しているとする。一方、エージェントに与えるタスクが決まれば、タスク達成には必要のないセンサも同時に決まる。このセンサを不用センサと呼び、不用センサを持つエージェントは過剰にセンサを所持していると言える。本論文では強化学習を対象として調査を行うことから、必要センサと不用センサを次のように定義する。

- ・必要センサ：タスクに与えられた報酬に対して従属である値を出力するセンサ
- ・不用センサ：タスクに与えられた報酬に対して独立である値を出力するセンサ

3. センサの過不足が学習に及ぼす影響の検証方法

本論文では、センサ数の過不足が学習に及ぼす影響を調査する。そのために、センサ数の異なるエージェントに同じタスクを与え、同じ学習手法を用いて学習させ、各エージェントの学習効率を比較する。

初めにセンサ数の異なるエージェントを用意する。まず、必要センサを持たず、0~M個の不用センサを持つエージェントを用意する。同様に、1~N個のセンサを持ち、0~M個の不用センサを持つエージェントを用意する。このとき、センサの数以外の要素が学習に影響を及ぼさないように、センサの能力を均一化する必要がある。そのために各センサが出力する値Vの最大値 V_{\max} を共通の値にする。

次に、エージェントが用いる学習手法について述べる。学習手法の違いによって学習の影響が変化することを防ぐため、全てのエージェントは強化学習を用いて学習を行う。行動学習手法にはQ学習、行動選択手法にはsoftmax法を用いる。Q学習の式を(2)に、softmax法の式を(3)に示す。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (2)$$

ここで $Q(s_t, a_t)$ は状態 s_t における行動 a_t のQ値を示している。また、 α は学習率、 γ は割引率を表している。

$$\pi(s_t, a_t) = \frac{e^{Q(s_t, a_t)/\tau}}{\sum_{b=1}^n e^{Q(s_t, b)/\tau}} \quad (3)$$

ここで $\pi(s_t, a_t)$ は状態 s_t における行動 a_t の選択確率である。また、 τ は温度と呼ばれる正定数である。

次に、エージェントが行うタスクについて述べる。エージェントに与えるタスクはルート探索のタスクである。タスクは複数の状態を持ち、各状態は二つの遷移先を持つ。エージェントはN個の必要センサを用いることでタスク内のすべての状態を認識できる。また、エージェントは各状態において行動を選択することで、行動に対応した遷移先の状態へ確定的に遷移する。そして、遷移先の状態に応じ

エージェントは報酬を獲得する。エージェントの目的は、このタスクにおいて最も高い報酬を得られ続けるルートを探索することである。このとき、高い報酬を得られる状態ほど辿り着くのが困難であるように報酬を設定した。各状態における報酬設定の式を(4)に示す。

$$Rwd_i = e^{\beta \cdot D_i} \quad (4)$$

ここで D_{ii} は状態 i から遷移して状態 i へ再び辿り着くまでに経由する最小の状態数である。 D_{ii} が大きい状態ほど高い報酬が設定される。また、 β は \exp の値を調整するための D_{ii} の係数である。

このタスクに対して各エージェントは R ステップの学習を行う。このときランダムな要素によって特定のエージェントに有利な状況が生じるのを防ぐため、 E 個のタスクに対して学習を行わせる。そして各タスクで得られた獲得報酬の平均から各エージェントの学習効率の違いを比較する。

次に、各エージェントの学習効率の比較方法について述べる。今回用いるタスクでは、より多くの報酬を獲得しているエージェントほど学習効率が高いと言える。そこで各エージェントが各ステップで獲得した報酬の推移を比較する。これにより各エージェントの学習速度の違いや、学習収束後の学習効率の違いについて比較する。しかし、今回用いるタスクでは高い報酬を得るために報酬の低い状態を経由する必要があるため、1ステップ毎の獲得報酬が安定しない。そこで t ステップ目における直近の L ステップで得た報酬の平均の推移を用いて比較を行う。この式を(5)に示す。

$$SMA = \frac{r_t + r_{t-1} + \dots + r_{t-L+1}}{L} \quad (5)$$

ここで r_t はエージェントが t ステップ目で得た報酬の値である。

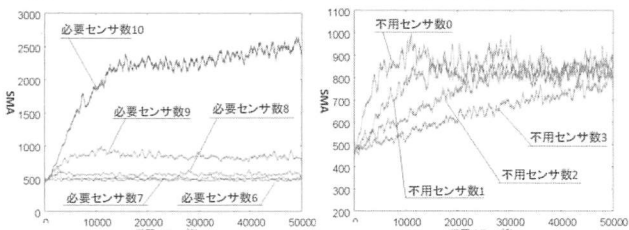
最後に、本実験で実際に用いたパラメータを表1に示す。

表1. 実験パラメータ

N(必要センサの最大数)	10	β	0.65
M(不用センサの最大数)	3	α	0.7
R(学習ステップ数)	50000	γ	0.6
E(実験を行うタスクの数)	20	τ	1000
V_{max} (センサの出力の最大値)	2	L	500

4. センサの過不足が学習に及ぼす影響の実験結果

初めに、エージェントが持つ必要センサ数が不足している場合の学習への影響を示す。不用センサ数0、必要センサ数6~10のエージェントのSMAの推移を図1(a)に示す。



(a) 必要センサ数6~10、不用センサ0のエージェントのSMAの推移 (b) 必要センサ数9、不用センサ0~3のエージェントのSMAの推移

図1. 各エージェントのSMAの推移の比較

図1(a)において、必要センサ数0~7のエージェントはSMAの推移の様子がほとんど同じであったので、必要センサ数5以下のエージェントのSMAについては省略した。各エージェントの学習収束後のSMAに注目すると、必要センサ数が減っていくにつれてSMAの値が小さくなっている。

このことから、タスクに対して必要センサが不足しているエージェントは学習収束後の学習効率が減少すると言える。

センサ数が不足しているエージェントは、本来切り分けるべき状態を一つの同じ状態として認識してしまう。つまり不適切な状態空間に対して学習を行うことになる。そのため、適切に学習を行っているにもかかわらず学習の対象が不適切であるため、学習効率が減少するのだと考えられる。

次に、エージェントが不用センサを持つ場合の学習への影響を示す。必要センサ数9、不用センサ数0~3のエージェントのSMAの推移を図1(b)に示す。各エージェントの学習収束のタイミングに注目すると、いずれのエージェントもSMAが900程度に到達した時点で学習が収束している。また、学習収束までのステップ数は不用センサの数に応じて増加している。今回は必要センサ数が9のエージェントの比較結果しか示していない。しかし必要センサ数が同じエージェント同士で不用センサの影響を比較した場合、不用センサの影響は全て同じ傾向であった。これらのことから不用センサを持つエージェントは学習収束に必要なステップ数が増加するといえる。また、学習収束後は不用センサを持たないエージェントと同じ学習効率を発揮できるといえる。

5. まとめ

本論文ではセンサ数の過不足が学習に及ぼす影響の調査を行った。調査は抽象化したセンサを対象とし、シミュレーションによって行った。実験ではセンサ数の異なるエージェントを複数用意し、各エージェントに同じ学習手法と同じタスクを与えて学習効率の違いを比較した。

調査の結果、エージェントが所持するセンサがタスクに対して不足すると学習収束後の学習効率が減少することを示した。また、エージェントが不用センサを所持すると、学習収束に必要なステップ数が増加することを示した。

初めに述べたように、現在のロボットの学習の研究ではセンサ数に過不足が生じたまま研究が行われている可能性がある。そこで本論文では、センサ数の過不足が学習に及ぼす様々な影響を調査した。その結果、センサ数に過不足が生じている場合、学習に様々な影響を及ぼすことを示した。そのため、センサの過不足が生じたまま学習手法の有用性の検証を行った場合に、学習手法が持つ本来の有用性が発揮できず、有用性の検証が適切に行えないと言える。以上のことから、ロボットの学習においては学習に合わせたセンサを搭載することが重要であることを示した。

今回の調査ではセンサ数の違いが学習に及ぼす影響に注目した。そのため各センサの能力は均一のものとして調査を行った。そこで今後の課題として、各センサの能力が異なる場合に学習に及ぼす影響について調査を行いたい。これによりロボットの学習におけるセンサの重要性をさらに示していきたい。

参考文献

- [1] Y. Onoue, K. Kurashige, "A relationship between ability of perception and learning efficiency", Proceeding of World Automation Congress 2008, 2008
- [2] 杉浦, 川上, 片井: "移動ロボットにおけるセンサ形態と制御系の同時設計法", 電学論C, vol. 128, no. 7, pp. 1-8, 2008.
- [3] 細田: "形態が学習にもたらすもの, 学習が形態にもたらすもの", 日本ロボット学会誌, vol. 22, no. 2, pp. 186-189, 2004.
- [4] R. S. Sutton, A. G. Barto, "Reinforcement learning", The MIT Press, 1998