

# 強化学習によるロボットの故障診断

藤巻 裕介<sup>†</sup> 中村 真吾<sup>‡</sup> 橋本 周司<sup>‡</sup>

早稲田大学先進理工学研究科<sup>†</sup> 早稲田大学理工学術院<sup>‡</sup>

## 1. はじめに

近い将来、ロボットは我々の生活に必要な不可欠な存在となると思われる。一般に、ロボットは事前にプログラムされた行動のみを行う。そのため、ロボットに故障が発生した場合には、作業を継続できないことが多い。一方、人間は怪我等によって身体の一部に不具合が生じた場合においても、その箇所を認識し、それを補う動きによって作業を継続することができる。このように、ロボットが自分自身で故障箇所を特定し、故障を補うような動作を行い、作業を継続することができれば、様々な状況において頑健性が高まることが期待される。従来、ロボットの故障診断では、故障パターンを予め用意することで対応してきたが、想定外の故障に対応することはできないといった問題がある。また、故障時のセンサ入力を学習データとしたニューラルネットワークを用いる手法[1]もあるが、学習データにない故障への対応は不可能であった。

本稿では、強化学習[2]を利用し故障診断を行う手法を提案する。強化学習は試行錯誤によって環境に適応する枠組みであり、ロボットが故障した場合であっても、その状況に適した動作を学習することができる。従来、強化学習はロボットの動作の学習[3]に用いられてきたが、故障診断には用いられた例はない。提案手法は、故障箇所の特定を強化学習による内部状態の変化から行うため、特別なセンサや事前の故障パターンに関する情報を必要としない。実験では、4足歩行ロボットのシミュレーションを行い、本手法の有効性を確かめた。

## 2. 提案手法

Fig.1 にロボットの学習の流れを示す。まず、正常時の動作を強化学習によって学習し、正常動作時に一定ステップで得られる報酬値を取得しておく。故障判定の際にはこの値を参照して、

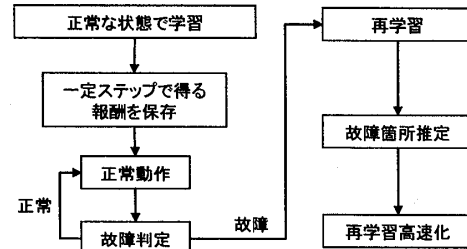


Fig.1 ロボットの動作フロー

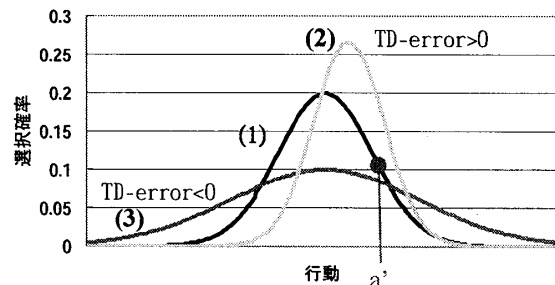


Fig.2 行動出力の分布変化

故障と判定した場合は再度強化学習を行い、得られた状態価値関数から故障箇所を推定する。故障箇所推定後は、推定結果を利用し故障を補う動作を獲得する。

### 2-1. 強化学習による目的動作の獲得

4足歩行ロボットの前進動作を獲得することを目的とした強化学習について述べる。各足にエンコーダ付のモータが1つずつ取り付けられ、各足の位置制御を行うことで前進移動を実現するロボットを考える。強化学習で扱う内部状態  $s = s(s_1, s_2, s_3, s_4)$  は、各足(1~4)の稼動域[-60,60](deg)を5分割し、合計で  $625(=5^4)$ 通りとした。また、行動  $a$  は、各足の次の目標角度への指令値とした。学習の報酬  $r$  は、単純に前に進む距離とし、状態  $s$  の価値を表す状態価値関数  $V(s)$  を TD-error 法を用いて更新する。また、各状態  $s_i$  は確率的に次の行動を決定するため、Fig.2 (1)のような正規分布を導入した。学習時には、TD-error 値が正の場合、つまりある状態  $s$  において選択した行動  $a'$  が良い結果であった場合は、Fig.2 (1)→(2)のように、状態  $s$  の正規分布の平均をその行動値

Fault diagnosis for Robots using Reinforcement Learning

<sup>†</sup> Yusuke Fujimaki, Graduate School of Advanced Science and Engineering, Waseda University

<sup>‡</sup> Shingo Nakamura, Shuji Hashimoto, Faculty of Science and Engineering, Waseda University

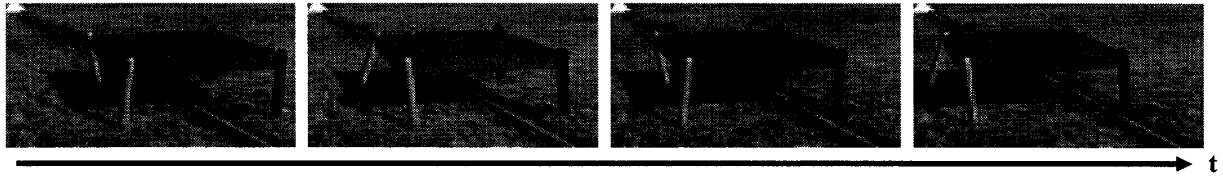


Fig.3 正常時におけるロボットの前進動作(左方向)

$a'$  に近づけ、分散を小さくする。反対に負の場合は、Fig.2 (1)→(3)のように分散の値だけを大きくする。

### 2-2. 故障判定と故障箇所推定

故障判定は、正常時の学習直後に  $n$  ステップ間で得られる総報酬  $R_0$  と、その後の  $n$  ステップ間で得られる総報酬  $R$  を比較することで行われる。本稿では、 $R$  が  $R_0$  の  $1/2$  未満であれば故障と判定することとした。故障箇所推定は、状態価値関数  $V(s)$  を用いて行う。例えば、もし足 1 が何らかの原因で故障したとすると状態価値関数は、 $s_1$  に依存しなくなるため、 $V(s_1, s_2, s_3, s_4) = V(s_2, s_3, s_4)$  となる。そこで、 $s_1$  に関する状態価値の分散値を  $B_{s_1}(s_2, s_3, s_4)$  とし、全ての状態の総和

$$F_{s_x} = \sum_{s_2} \sum_{s_3} \sum_{s_4} B_{s_1}(s_2, s_3, s_4) \quad (1)$$

を求める。これを他の足に関しても計算し、 $F_{s_x}$  が最小となる足  $x$  が故障箇所であるとする。

### 2-3. 故障箇所推定後の学習高速化

強化学習は状態空間が大きいほど学習に時間を要する。故障時における動作は、故障した箇所に依存しないと考えられるため、状態空間を小さくすることで学習の高速化が期待できる。具体的には、足 1 が故障したとすれば内部状態  $s$  を  $s=(s_2, s_3, s_4)$  と定義しなおし、 $s_1$  の状態については考慮しない。これによりロボットの状態数を  $625(=5^4)$  通りから  $125(=5^3)$  通りに減らすことができる。

## 3. 実験

まず、正常時のロボットに一定回数の学習をさせ、前進動作を獲得させた。なお、ロボットのサイズは  $300 \times 150 \times 120$  [mm] とした。このときのロボットの動作の様子を Fig.3 に示す。次に故障が発生したとして、左後の足 4 をモータから取り外し、学習を行った。

Fig.4 に故障後の学習結果から得られた  $F_{s_x}$  を示す。故障した左後足の場合に  $F_{s_x}$  が小さくなっており、故障箇所の特定が可能であることがわかる。

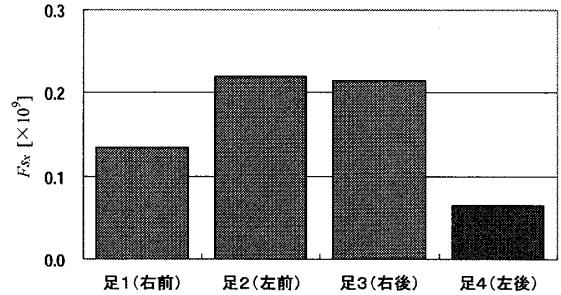


Fig.4 故障時の  $F_{s_x}$

次に提案手法 2-3 の有効性を検証するため、足 4 を外したロボットで、状態を  $s = s(s_1, s_2, s_3, s_4)$  と定義した場合、及び  $s = s(s_1, s_2, s_3)$  と定義した場合について、速度が  $10\text{cm/min}$  を超えるまでの学習回数を調べた。前者が  $170,000$  回、後者が  $50,000$  回であり、提案手法 2-3 を用いることによって学習の高速化が可能であることが判る。

## 4. おわりに

強化学習を用いることでロボット自身が歩行動作を獲得できるばかりでなく、学習によって得た状態価値関数を解析することで、故障箇所を推定することが可能であることを示した。さらに推定結果を利用し故障後の学習を高速化できることを示した。今後の課題としては、本手法の適用可能範囲の検討などある。

### 謝辞

本研究の一部は、早稲田大学ヒューマノイド研究所、グローバル COE プログラム「グローバル ロボット アカデミア」、科学技術振興機構 CREST 研究「人を引き込む身体的メディア場の生成・制御技術」の研究助成を受けて行われた。

### 参考文献

- [1] S. ZHANG, T. ASAKURA, X. XU, B. XU, "Fault Diagnosis System for Rotary Machine Based on Fuzzy Neural Networks," JSME international journal. Series C, Mechanical systems, machine elements and manufacturing, Vol.46, No.3(20030915) pp. 1035-1041.
- [2] R. S. Sutton, A. G. Barto, Reinforcement Learning: An Introduction. Cambridge, MA: MIT Press, 1998.
- [3] H. Kimura, T. Yamashita, S. Kobayashi, "Reinforcement Learning of Walking Behavior for a Four-Legged Robot," in Proceedings of the 40th IEEE Conference on Decision and Control, Orland, Florida USA, December 2001.