

# 安定性の評価とグラフ埋め込みによるトラスの施工経路の強化学習 Reinforcement Learning for Construction Process of Truss by Stability Assessment and Graph Embedding

○小寺 正也\*<sup>1</sup>, 林 和希\*<sup>2</sup>, 大崎 純\*<sup>3</sup>

Masaya KOTERA \*<sup>1</sup>, Kazuki HAYASHI \*<sup>2</sup> and Makoto OHSAKI \*<sup>3</sup>

\*1 京都大学工学研究科建築学専攻 修士課程

Graduated Student, Department of Architecture and Architectural Engineering, University of Kyoto.

\*2 京都大学工学研究科建築学専攻 助教 博士(工学)

Assistant Professor, Department of Architecture and Architectural Engineering, Kyoto University, Ph.D.

\*3 京都大学工学研究科建築学専攻 教授 博士(工学)

Professor, Department of Architecture and Architectural Engineering, Kyoto University, Ph.D.

キーワード：施工経路最適化；トラス；強化学習；グラフ埋め込み

Keywords: Construction sequence optimization; truss; reinforcement learning; graph embedding.

## 1. はじめに

建築構造における最適化問題は多岐にわたり、例えば施工経路最適化 [1], トポロジー最適化 [2], 部材配置最適化 [3] が挙げられる。中でも施工経路は施工時の安全性や効率性を考えるうえで重要な要素であり、過去には遺伝的アルゴリズムや動的計画法などを用いて一般的なトラス構造物やケーブル構造に対して目的に適った施工経路を獲得する方法が提案されている [1, 4]。ここでの目的は、安全性と施工効率性の観点から問題ごとに設定される。既往研究 [1, 4] においては、数値解析による手法を用いて部材に作用する応力や仮設支持点での反力に上限を与え、施工過程中的トラスやケーブル構造がその制約を満足することが目的とされていた。本研究では、とくにトラスの安定性に着目し、施工過程中的トラスの不安定次数が最小となることを目的とすることで、仮設支持数を抑えた低コストかつ効率的な施工経路を生成するための強化学習に基づく手法を提案する。

状態の組み合わせが膨大な強化学習タスクに対してはニューラルネットワークなどの関数近似を用いた縮約表現がしばしば用いられる。ただし、トラスは多くの部材が複雑な接続関係をなして力を伝達しているため、トラスを対象として関数近似を用いた強化学習を行う場合、トラスの接続関係を関数近似過程でどのように考慮するかが大きな課題となる。本研究では、トラスの接続関係を考慮した各部材の定量的な情報を抽出するため、ウェブサイト間のリンク予測やたんばく質の解析に用いられるグラフ埋め込みを導入する。

グラフ埋め込みはエッジとノードで構成されたグラフ構造をベクトル空間に埋め込む手法の総称であり、本研究においては平面的なトラスモデルを部材(エッジ)と節点(ノ

ード) からなるグラフ構造とみなし、トラスの構造性状を各部材の特徴ベクトルとして抽出する。グラフ埋め込みを用いれば、近傍の部材や節点の数値情報だけでなくその接続関係を考慮できるというメリットがある。構造物の安定性は構成要素である部材や節点の相対的な位置関係に大きく影響されるため、接続関係を考慮できるグラフ埋め込みにより効果的にトラスの安定性を推定できると考えられる。なお、トラスの最適設計に強化学習とグラフ埋め込みを組み合わせた手法を適用した例としては、密に接続したグラウンドストラクチャから逐次不要部材を除去して疎なトポロジーを得るトポロジー最適化がある [2]。

本研究の目的は、安定性を確保した施工経路を生成するために強化学習とグラフ埋め込みを併用 [5] する際に必要な報酬等の諸設定を提示することである。また、その過程で計算負荷を軽減するため構造解析を経ずに構造物の幾何的な特徴を抽出対象とし、学習結果を学習対象でないトラスモデルにも適用できる汎用性を獲得できることを示す。

## 2. Q-learning とグラフ埋め込みの概要

トラス構造物のモデルとして、図 1 のような平面トラスを考える。

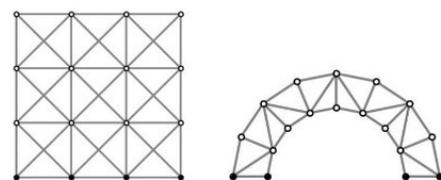


図 1 学習対象モデル

学習を行う上で、完成状態からすべての部材が除去されるまで一つずつ部材を除去することを 1 エピソードとし、施工経路はその逆の手順として与えられると考える。1 エピソード内で部材を除去した数を時刻 $t$ 、時刻 $t$ におけるトラスの特性を状態 $s_t$ 、時刻 $t$ において除去する部材を選択することを行動 $a_t$ と定義すると、ある状態 $s_t$ から行動 $a_t$ を決定した際に次の状態 $s_{t+1}$ は唯一に定まることから、この過程はマルコフ決定過程である。

ここで、ある状態 $s_t$ から行動 $a_t$ を決定した際に、その行動 $a_t$ における目的の満足度を定量的に表す指標として、報酬 $R_{s_t, a_t}$ を定義する。報酬 $R_{s_t, a_t}$ は部材を除去するごとに与えられ、また、ある状態 $s_t$ から行動 $a_t$ を決定した後に得られることが見込まれる割引報酬和をその決定の行動価値 $Q_{(s_t, a_t)}$ と定義する。このとき、以下のベルマン方程式が成立する。

$$Q_{(s_t, a_t)} = R_{s_t, a_t} + \gamma \cdot \max_{a_{t+1}} (Q_{(s_{t+1}, a_{t+1})}) \quad (1)$$

したがって、ベルマン方程式(1)の両辺が等しくなるように行動価値 $Q_{(s_t, a_t)}$ を推定することが望ましく、両辺の誤差を最小化することがQ-learningの目標となる。Q-learningにより正しい行動価値の値を推定できるようになれば、各状態において最も行動価値の高い行動を選択することで最適な逆施工経路が与えられる。

ここで、行動価値 $Q_{(s_t, a_t)}$ を直接学習する場合、状態数×行動数の膨大な組み合わせ数の $Q_{(s_t, a_t)}$ を推定する必要があり、現実的でない。この組合せ爆発による問題を解決するため、状態を縮約してトラスの安定性に関する情報をグラフ埋め込みを用いて表現する。図2にグラフ埋め込みの概要を示す。

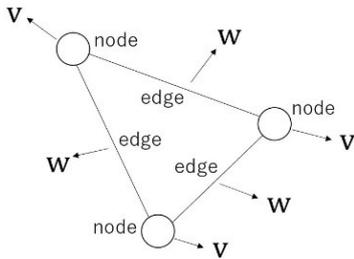


図2 グラフ埋め込みに用いる節点入力 $v$ と部材入力 $w$

$w$ の次元を $n_w$ 、 $v$ の次元を $n_v$ とおき、学習可能な重み係数 $\theta_1 \in \mathbb{R}^{n_r \times n_w}$ 、 $\theta_2 \in \mathbb{R}^{n_r \times n_r}$ 、 $\theta_3 \in \mathbb{R}^{n_r \times n_v}$ 、 $\theta_4 \in \mathbb{R}^{n_r \times n_r}$ 、 $\theta_5 \in \mathbb{R}^{n_r \times n_r}$ 、 $\theta_6 \in \mathbb{R}^{n_r \times n_r}$ を用いて、部材を除去する際に式(2)、(3)のように部材の特徴量 $\mu$ を更新する。 $n_r$ は特徴量 $\mu$ の次元を表し、以下の例では100とする。また、同様に学習可能な重み係数 $\theta_7 \in \mathbb{R}^{2n_r \times 1}$ 、 $\theta_8 \in \mathbb{R}^{n_r \times n_r}$ 、 $\theta_9 \in \mathbb{R}^{n_r \times n_v}$ と更

新した $\mu$ を用いて行動価値を算出する。

$$\mu_i^{(0)} = 0 \quad (2a)$$

$$\mu_i^{(t+1)} = \text{ReLU}(h_1^{(t)} + h_2^{(t)} + h_3^{(t)} + h_4^{(t)}) \quad (2b)$$

$$h_1^{(t)} = \theta_1 w_i^{(t)} \quad (2c)$$

$$h_2^{(t)} = \theta_2 \sum_{j=1}^2 \text{ReLU}(\theta_3 v_{i,j}^{(t)}) \quad (2d)$$

$$h_3^{(t)} = \theta_4 \mu_i^{(t)} \quad (2e)$$

$$h_4^{(t)} = \theta_5 \sum_{j=1}^2 \text{ReLU}\left(\theta_6 \sum_{k \in \Phi_{i,j}} \mu_k^{(t)}\right) \quad (2f)$$

$$Q_{(\hat{\mu}, i)} = \theta_7^T \left( \text{ReLU} \left[ \theta_8 \sum_{i=1}^{n_m} \mu_i^{(T)} ; \theta_9 \mu_i^{(T)} \right] \right) \quad (3)$$

$\theta$ が学習されることによって間接的に行動価値が学習され、学習が理想的に進めばベルマン方程式の両辺の差が0になる値へ収束するため、差の二乗を損失関数としそれを最小化する。更新量はRMSpropによって損失関数の勾配の大きさに基づいて決定される。学習は以下のように定式化される。

$$\text{minimize} \left( R_{s_t, a_t} + \gamma \max_{a_{t+1}} Q_{(s_{t+1}, a_{t+1}, \bar{\theta})} - Q_{(s_t, a_t, \theta)} \right)^2 \quad (4)$$

$$(\theta = \{\theta_1, \theta_2, \dots, \theta_9\})$$

ここで、 $\gamma$ は割引率であり、直近の報酬と後に得られる報酬の優先度を定める。以下の例では $\gamma = 0.99$ とし、遅延報酬を考慮できるように設定している。また、 $\bar{\theta}$ は学習過程で得られた $\theta$ の過去の値であり、 $\bar{\theta}$ を用いて式(4)の第三項の $\theta$ を更新することで学習を安定化させる[6]。

### 3. 報酬、入力と学習方法

施工経路を目的に沿うように正しく学習するためには、報酬、入力を適切に設定する必要がある。

まず、除去後の状態における不安定次数が大きいほど小さい値が与えられるように報酬を設定する。また、実際は支点から順に施工が行われるため、逆施工経路としては支点から離れている部材が優先的に除去されるのが望ましい。したがって、支点からの距離が大きいほど報酬が大きく与えられるように、次式で報酬を設定する。

$$R_{s_t, a_t} = \frac{1}{2^{J_{s_{t+1}}}} - \frac{1}{1 + e^{\frac{l_i + 4 - \max_{k \in \Phi_{s_t}} l_k}{2}}} \quad (5)$$

$R_{s_t, a_t}$ : 状態 $s_t$ で行動 $a_t$ を選んだ際の報酬

$J_{s_t}$ : 状態 $s_t$ における不安定次数

$l_i$ : 除去対象のトラス部材*i*の支点からの最短距離  
 $\Phi_{s_t}$ : 状態 $s_t$ で残っている部材の集合

著者らの以前の研究[7]では報酬を式(6)で定義しており、それに比べて式(5)では不安定次数が小さい場合の報酬差をより大きく、支点からの距離が大きい場合の報酬差をより小さくしている。本報告では、式(5)を用いた際の結果と式(6)を用いた際の結果を比較する。なお、 $l_{ave}$ は平均部材長、 $F$ は現状態におけるトラスの不安定次数である。

$$R_{s_t, a_t} = \left( \frac{l_i}{\max_{k \in \Phi_{s_t}} l_k} \right)^m - \left( \frac{J_{s_{t+1}}}{F} \right)^n \quad (6a)$$

$$m = \begin{cases} \frac{\ln 0.8}{\ln \left( 1 - \frac{l_{ave}}{\max_{k \in \Phi_{s_t}} l_k} \right)} & \dots (\max_{k \in \Phi_{s_t}} l_k \geq 5l_{ave}) \\ 1 & \dots (\max_{k \in \Phi_{s_t}} l_k < 5l_{ave}) \end{cases} \quad (6b)$$

$$n = - \frac{\ln \left( 1 - \frac{\ln 2}{\ln F} \right)}{\ln 2} \quad (6c)$$

ここで、行動価値は報酬によって決定され、行動価値を学習するために必要な入力是不安定次数と支点からの距離の二つの指標とする。したがって、まず部材の入力に関して、部材の重心と支点からの最短距離と、部材の除去情報の二つの情報を表1のように設定する。また、節点の入力は、支点からの最短距離と接続している部材数として表2のように設定する。

表1: 部材*i*の入力 $w_i$

Index	入力内容	式
1	支点からの最短距離	$\frac{l_i}{\max_k l_k}$
2	除去されているか否か	$\begin{cases} \text{除去済み: 1} \\ \text{未除去: 0} \end{cases}$

表2: 節点*k*の入力 $v_i$

Index	入力内容	式
1	支点からの最短距離	$\frac{l_i}{\max_k l_k}$
2	接続部材数	$\frac{x}{4^{x-1}}$

部材と節点の両方に支点からの最短距離を入力しているが、これは強化学習過程で支点から離れた部材を優先的に選択するために必要である。また、隣接部材が少ないほど、その部材を除去した際に不安定な構造が生じやすいと考

えられるため、部材の除去に関するフラグを部材の入力に、接続部材数を節点の入力に付加している。

これらの情報を用いて、4000 エピソード繰り返して学習する。ただし、汎用性の高い学習結果を得るためには一つのトラスモデルでの学習だけでは不十分であり、学習結果を学習対象のモデルと形状が大きく異なるモデルに適用すると安定性が確保されないなどの問題がしばしば生じる。これを解決するため、学習対象のモデルをエピソードごとに変更しながら学習を進める。

学習過程で重み $\Theta$ が逐次更新されるが、10 エピソードごとに現在の $\Theta$ を用いてシミュレーションを行い、過去のシミュレーションで得られた累積報酬以上の累積報酬を獲得できれば、その時点での $\Theta$ を最も良い性能を発揮する学習パラメータとして記録する。ただし、二種類のトラスを用いて学習する場合にはそれぞれのトラスに対するエージェントの性能を同時に評価しなければならない。そのため、各学習対象のトラスに対してシミュレーションを行い得られた累積報酬の積を用いて $\Theta$ を評価する。

#### 4. 学習結果

式(5)の報酬式を用いて図1に示す二つのモデルに関して交互に学習したときの学習経過を図3に示す。グラフの横軸は学習回数、縦軸はその時点での各学習対象についてシミュレーションを行い得られた累積報酬和の積を表している。学習エピソードが増加するにつれて、得られる累積報酬が大きくなり、収束していることが読み取れる。

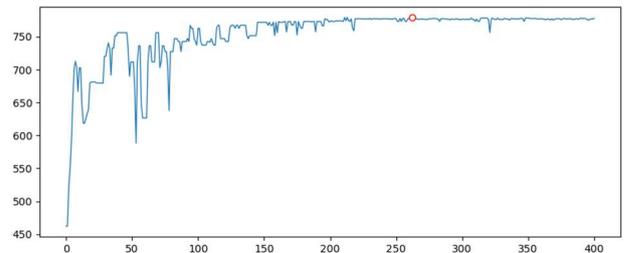


図3 学習経過 1

最適化された学習パラメータ $\Theta$ をそれぞれのモデルに適用すると、図4のような施工経路が得られた。いずれのモデルに対しても不安定次数が可能な限り小さくなるよう、支点から離れた部材から除去されていることが確認でき、目的に適った施工経路が得られた。

式(6)の報酬式を用いて学習した場合の学習経過を図6に示す。エージェントの性能が収束するまでの学習回数は式(5)の報酬式を用いた場合と比較してそれほど差がないことが読み取れる。一方で、学習結果を式(6)を用いた場合と同じ学習対象外トラスに適用した結果、図7のような不安定次数の高い状態が生じて安定性に劣る経路が導か

れてしまう。

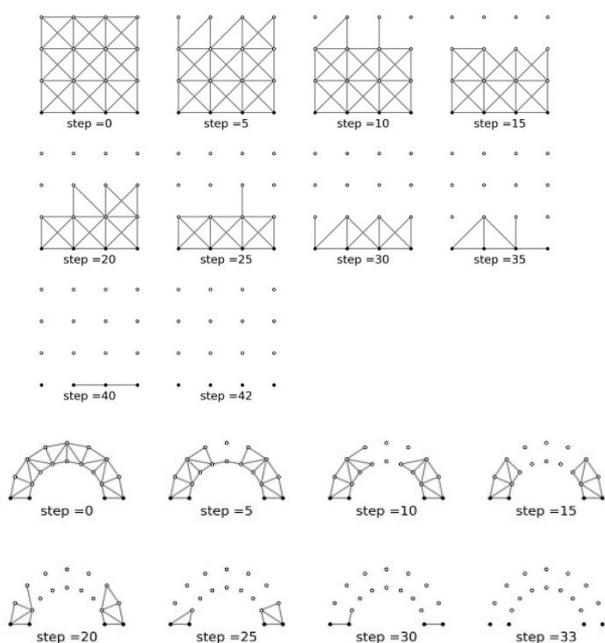


図4 学習対象トラスの部材除去過程(逆過程が施工経路)

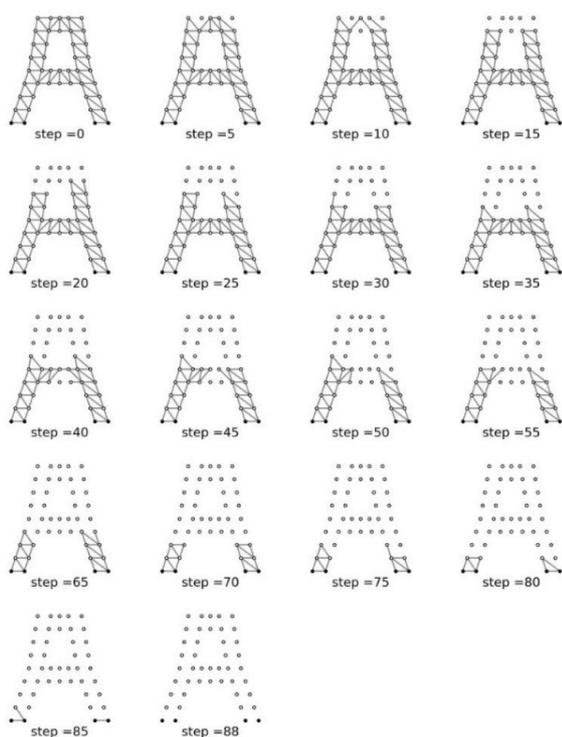


図5 学習に用いなかったトラスの部材除去過程(式(5)を用いて学習した場合)

## 5. 結論

強化学習とグラフ埋め込みを併用し、安定性を考慮した施工経路を求める手法を提案した。その過程で構造解析を用いず、未学習のモデルに対しても適用できるような報酬、入力と学習方法を提案した。とくに、著者らの以前の研究

から、不安定次数が小さい場合の報酬差が大きく、支点からの距離が大きい場合の報酬差が小さくなるように報酬設定を変更したことにより、学習の速度や収束性を損わずに学習結果の汎用性の改善を実現できた。

今後の課題としては、さらに広範囲のトラスモデルに適用できる汎用性の獲得や、対象を3次元トラスに拡張したうえで同様の学習結果を実現できる報酬、入力と学習方法を提案することが考えられる。

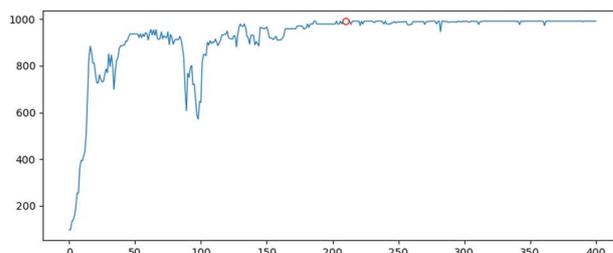


図6 学習経過 2

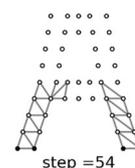


図7 学習に用いなかったトラスの部材除去過程(式(6)を用いて学習した場合)

## 謝辞

本研究は、JSPS 科研費 (JP20H4467, JP21K20461) の助成を得た。

## [参考文献]

- 金子義信, 兼光知己, 三井和男, 登坂宣好, 半谷裕彦, 遺伝的アルゴリズムによるトラス構造物の施工経路最適化, 日本建築学会構造系論文集, Vol. 63(508), pp.87-92, 1998.
- K. Hayashi and M. Ohsaki, Reinforcement learning and graph embedding for binary truss topology optimization under stress and displacement constraints, Frontiers in Built Environment, Vol. 6, Paper No. 59, 2021.
- T. Tamura, M. Ohsaki and J. Takagi, Machine learning for combinatorial optimization of brace placement of steel frames, Japan Architectural Review, Vol. 1(4), pp. 419-430, 2018.
- 藤原淳, 大崎純, 水谷太朗, 北折智規, 加藤直樹, 細澤 治, ケーブル補強骨組構造物の張力および施工順序最適化, 日本建築学会構造系論文集, Vol. 67(556), pp. 101-107, 2002.
- H. Dai, E. B. Khalil, Y. Zhang, B. Dilkina and L. Song, Learning combinatorial optimization algorithms over graphs, Proc. 31st Int. Conf. on Neural Information Processing Systems (NIPS 2017), pp. 6351-6361, 2017.
- V. Mnih et al., Human-level control through deep reinforcement learning, Nature, Vol. 518, pp. 529-533, 2015.
- 小寺正也, 大崎純, 林和希, 強化学習とグラフ埋め込みによるトラスの安定な施工経路生成, 日本建築学会大会学術講演梗概集, 構造 I, pp. 777-778, 2021.