

# 強化学習による最小費用流問題の ヒューリスティック解に関する検討

岩本 圭介<sup>†</sup> 福田 龍樹<sup>†</sup>

<sup>†</sup> 北九州工業高等専門学校 生産デザイン工学科

## 1. はじめに

従来より、最短経路問題や巡回セールスマン問題などを筆頭とする組み合わせ最適化問題の最適解およびヒューリスティック解を求めるために、グラフ理論を用いたアルゴリズムが提案されてきた。しかし、近年の情報化社会において情報流通やソーシャルネットワークなどがより複雑さを増しているため、求解の更なる高速化が必要とされている。そこで、本研究では計算時間の削減を目的とする、組み合わせ最適化問題の一つである最小費用流問題のヒューリスティック解を強化学習を用いて解く方法について考察する。

## 2. 最小費用流問題

頂点の集合が $V$ 、辺の集合が $E$ である重み付き有向グラフ $G = (V, E)$ において、二つの頂点 $u, v$ を結ぶ辺を $e = (u, v) \in E$ としたときの各辺 $e$ の容量を $c(e)$ 、コストを $d(e)$ とする。始点から終点に流量 $F$ のフローを流したときの、各辺 $e$ に流れるフロー $f(e)$ とコスト $d(e)$ の積の総和 $s$ は(1)式で表される[1]。

$$s = \sum_e \{f(e) \times d(e)\} \quad (1)$$

最小費用流問題とは、 $s$ が最小となるフロー $f(e)$ を求める問題であり、ダイクストラ法などを用いることで時間計算量 $O(F|E| \log|V|)$ または $O(F|V|^2)$ で最小費用流を求めることが可能である[1]。

## 3. Pointer Networks

Sequence-to-sequence モデルとは、入力配列からエンコーダのRNN (Recurrent Neural Network)を用いてベクトルを生成し、そのベクトルからデコーダのRNNを用いて出力配列を生成する機構である[2]。しかし、このモデルでは出力の次元数が固定されるため、最小費用流問題に関してはテストデータの次元数がトレーニングデータの次元数に依存する問題がある[3]。そこで、本研究では次元数固定の制約に対処するために Pointer Networks モデルを用いる。Pointer Networks は、Sequence-to-sequence と同様にエンコーダとデコーダの二つのRNNを持ち、その両方がLSTM (Long short-term memory)で構成されている[4]。このモデルでは、入力配列の特定の位置を指すポインタを出力する Pointing 機構を用いることで次元数が可変な場合に対応する[2]。Pointer Networks の構造を図1に示す[2]。

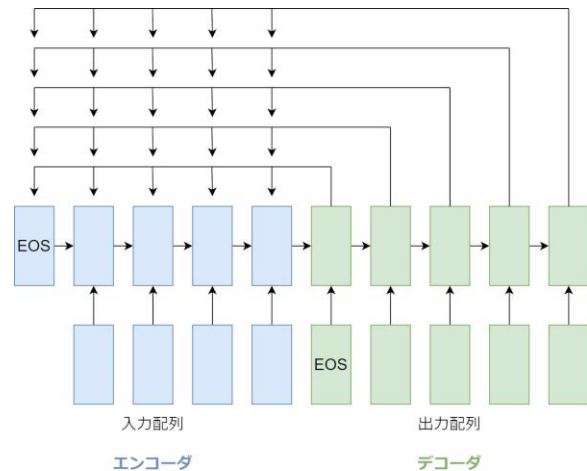


図1. Pointer Networks の構造

## 4. 強化学習アルゴリズム

今回は強化学習アルゴリズムに Actor-Critic を用いる。Actor-Critic とは、Actor (行動) と Critic (評価) の両方を学習するアルゴリズムであり、Actor のみの場合よりも収束が早くなることが期待される[5]。

## 5. 今後の研究

これまでに、最小費用流問題の理解および組み合わせ最適化問題に強化学習を用いる手法の調査を行った。

今後の展望として、Pointer Networks を用いた最小費用流問題のヒューリスティック解を求めるプログラムの構築を行う。その後、構築したプログラムの計算時間およびこの手法で求めたヒューリスティック解の精度について調査を行い考察する。また、強化学習アルゴリズムに Actor-Critic 以外を用いるとどうなるか検証していく。

## 参考文献

- [1] 秋葉 拓哉, 岩田 陽一, 北川 宜稔, “問題解決のアルゴリズム活用力とコーディングテクニックを鍛える プログラミングコンテストチャレンジブック 第2版”, マイナビ出版, pp. 188-224, 2012.
- [2] Ilya Sutskever, Oriol Vinyals, Quoc V. Le, “Sequence to Sequence Learning with Neural Networks”, arXiv, 2014.
- [3] Oriol Vinyals, Meire Fortunato, Navdeep Jaitly, “Pointer Networks”, arXiv, 2017.
- [4] Irwan Bello, Hieu Pham, Quoc V. Le, Mohammad Norouzi, Samy Bengio, “NEURAL COMBINATORIAL OPTIMIZATION WITH REINFORCEMENT LEARNING”, ICLR 2017, 2017.
- [5] Vijay R. Konda, John N. Tsitsiklis, “Actor-Critic Algorithms”, NeurIPS2000, pp. 1008-1014, 2000.