

鉄道運転整理における強化学習を用いたリスケジューリングの検討 A Study of Rescheduling with Reinforcement Learning for Railway Operation Control

福島 卓弥[†] 松本 杜青[†] 赤塚 駿一[†] 高田 晋太郎[†] 寺本 やえみ[†]
Takuya Fukushima Morio Matsumoto Shunichi Akatsuka Shintaro Takada Yaemi Teramoto

1. はじめに

複雑化した社会システムの運営には、綿密な計画が重要である。一方で、計画が不測の事態によって遂行できなくなった際は、遂行可能な状態に計画を修正するリスケジューリングを行う必要がある。リスケジューリングは、制約を満たしながら、複数の修正操作を組み合わせることで短時間で再計画する必要があり、人間の負担が大きく自動化が求められている。

計画立案に有用な手法として、強化学習が注目を集めている。強化学習は、複数の計画操作に対して、時系列を考慮して行動を決定することに優れ、将来の獲得報酬を最大化するような行動をとる方策をあらかじめ学習させることで、適用時にはその学習済みの方策を効率的な解探索に活用できるため、短時間で計画立案が可能である。

本論文では、強化学習をリスケジューリング問題に適用する。リスケジューリングを行う問題設定をタスクとし、タスクごとにリスケジューリングを考える。リスケジューリング問題は、ナース勤務表や生産管理表の計画修正など様々であるが、特に、本論文では、多様なタスクパターンがあり、かつ複数の修正操作を組み合わせる必要のあるリスケジューリング問題である「鉄道運転整理」に着目した。鉄道運転整理に対して強化学習を適用し得られた計画を、複数の行動候補から適切なものを選択可能か評価することで、リスケジューリング問題に対する強化学習の有用性を検証する。

2. 鉄道運転整理

鉄道運転整理とは、当日の運行に乱れが生じたとき、旅客への影響や乗務員への負担、列車の走行速度など、輸送システム全体を考えた多目的、大局的な判断を必要とする運転計画のリスケジューリング作業のことである[1]。本論文では、ある支障に対する一連の運転整理の集合を「運転整理案」と呼ぶ。運行乱れが生じてから運転整理案を立案するまでの時間は限られているが、近年は、列車運行の複雑化が進んでいることから、運転整理業務は難しく時間のかかる作業になっている。

列車の運行には制約が多く存在し、それらの制約を満たされるように運転整理を行う必要がある。制約については、[1]を参照されたい。実行される運転整理は支障によって異なるが、一般的に用いられる運転整理を以下に列挙する[2]。

- 運休：指定した区間において、指定した列車の運行を取りやめる。特に、運休区間に始発駅終点駅を含まない運休を運休①、運休区間の片側に始発駅または終点駅を含む運休を運休②、始発駅から終点駅までを全て運休することを運休③とする。
- 特発：指定した駅において、指定した列車の走行を別の列車に割り当てる。

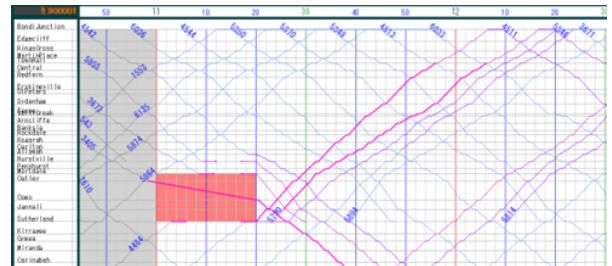


図 1 シドニーT4線(4章で後述)で支障(赤い領域)が生じた際のダイヤの例。予定時刻通りに走行可能な列車は青、遅延が生じている列車はピンクで表されている。縦軸は駅、横軸は時間を示している。

3. 提案手法

本論文では、遅延の生じている列車を運転整理対象として、各列車に対して予定出発時刻順に行動を決定しながら状態を遷移するシングルエージェント環境の強化学習を適用することを提案する。エージェントの行動として、運転整理のうち運休(運休①、運休②、運休③)と特発の4種類を割り当て、加えて「運転整理を行わない」を含めた計5種類を行動候補とする。エージェントが観測する状態として、現在の総遅延時間や該当列車の遅延時間といった遅延に関する情報、上りや下りなどの走行方向などの列車に関する情報、走行駅に関する情報を与える。報酬関数は、遅延を解消することによる報酬と、運休や不十分なダイヤに対するペナルティを付与して定める。運転整理対象の各列車に対して運転整理を終えるまでを1エピソードとし、1エピソード内における、ある支障に対するリスケジューリング完遂時の獲得報酬が高い一連の運転整理を、強化学習を用いることで探索する。

4. 数値実験

4章では、提案手法を数値実験により評価する。

4.1 線路データ

本実験では、提案手法の有用性を確認するため、オープンデータ[3]であるシドニー路線(RAILTEX)のT4-Lineを対象として実験を行う。T4-Lineは、単線区間がなく、特急列車もないため、運転整理を行う路線の中では簡単な路線である。T4-Lineにて支障が生じた際の例を図1に示す。T4-Lineの駅のうち、発着ホームが4か所以上ある駅にて列車の折り返し運転が可能であると仮定し、折り返し可能駅間を運休するように環境を設計した。T4-Lineでは、運転整理のうち、運休①もしくは特発を適切な箇所で行うことで、獲得報酬値の高い運転整理案が導かれる。そのため、強化学習は、5種類の行動から適切な箇所で行う行動(運休①、特発、整理しない)を選択する必要がある。

[†]日立製作所 研究開発グループ
Research & Department Group, Hitachi, Ltd.

4.2 支障パラメータ

支障は以下の 3 つのパラメータで定まると仮定する。

- 支障駅：支障駅が起きている区間。本論文では、折り返し可能駅のうち指定された 2 駅間で支障が発生すると仮定する。
- 支障開始時刻：支障が始まる時刻。
- 支障時間：支障が継続する時間。本報告では、支障が何分間続くのかを過去の経験から見積もることができると仮定する。

本論文では、支障開始時刻を 11 時に固定し、2 パターンの支障に対して、強化学習を適用した。始発駅と終点駅を支障駅に含めない支障設定を採用し、支障駅間は Mortdale から Sutherland の 5 駅とした。支障時間は以下の 2 パターンにて検証した。

- I. 支障時間 短 (20 分)
- II. 支障時間 長 (60 分)

図 1 はパターン I の運転整理前のダイヤである。

4.3 学習設定

強化学習アルゴリズムとして、深層強化学習の一種である Rainbow [4] を採用した。探索戦略として Annealing ϵ -greedy を用いた。学習は各支障に対して 1000 エピソード行い、学習中に得られた獲得報酬が最大となる 1 エピソード内の行動列を強化学習の解 (運転整理案) とする。

4.4 評価方法

運転整理の評価方法として、強化学習で立案した運転整理案と、強化学習を用いずにヒューリスティックなルールと探索を用いて作成した運転整理案の遅延解消時間との比較によって評価する。ヒューリスティックなルールは、以下のように定める。

- 支障発生時に支障駅間を走行している列車に対して、特発を実行
- 支障により遅延が生じている上記以外の列車に対して、遅延時間の大きい列車から順に n 本を運休
- 最も報酬関数が高くなる n を総当たりで探索 ($0 \leq n \leq N_{train}$, N_{train} : 遅延が生じている列車数)

上述のルールは遅延時間の解消をめざしたルールになっている。本実験で用いる単線区間や特急列車のないシドニー路線に対しては、遅延時間を解消する運転整理案を立案可能であるが、その他の複雑な路線に対して有用とは限らない点に留意する。上述のルールによる解と強化学習の解を比較するため、以下の式で運転整理案を評価する。

$$(\text{Score}) = (\text{DelayTime}_{\text{RL}}) / (\text{DelayTime}_{\text{rule}})$$

ここで、 $\text{DelayTime}_{\text{RL}}$ は強化学習による運転整理後の総遅延時間を、 $\text{DelayTime}_{\text{rule}}$ はルールによる運転整理後の総遅延時間を表す。遅延解消時間は、運転整理前後における各列車の総遅延時間の差によって算出する。

4.5 実験結果・考察

強化学習によって得られたパターン I に対する運転整理後の運行ダイヤを図 2 に示す。運転整理を行わない場合のダイヤである図 1 における各列車の総遅延時間は約 2037 分であった。強化学習によって運転整理が行われ、運休①を 6 回、特発 1 回実行した結果、総遅延時間を約 38 分に減少させることを確認した。

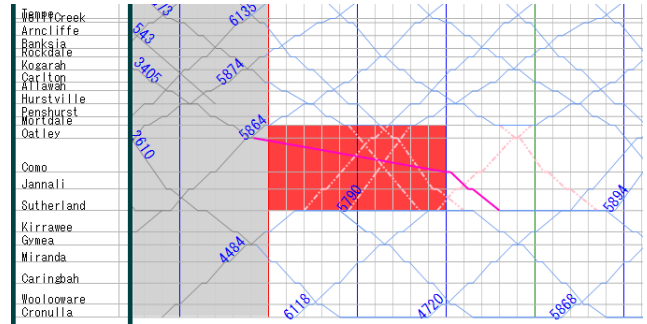


図 2 支障パターン I に対する強化学習による運転整理後の運行ダイヤ。

表 1 各支障に対する Score と、強化学習とルールによる運転整理後の総遅延時間。

Pattern	Score	DelayTime _{RL}	DelayTime _{rule}
I	1.00	38 mins.	38 mins.
II	1.00	162 mins.	162 mins.

各支障パターンに対する Score や総遅延時間を表 1 に示す。各支障パターンでの Score が 1.0 であることから、強化学習がルールと同じ運転整理案を見つけ、遅延時間が 98% 以上解消されたことを確認した。このことから、強化学習が複数の行動選択肢から有用な運転整理案を探索できていると分かる。

実際の路線では、単線区間や特急列車などで運転整理が複雑になり、遅延の大きい列車から順に運休するルールでは対応できない場合が多く存在する。強化学習は、解空間の中から有用な運転整理案を探索できることが示されたため、ルールで対応できない場合においても適切な行動を探し得ると期待できる。

様々な支障パターンにおける鉄道運転整理に対して、適切な行動を選択肢できることが分かった。以上より、複数の行動選択肢があるリスケジューリング問題において、強化学習が有用であるという見込みを得た。

5. おわりに

本論文では、多様なタスクパターンがあり、かつ複数の修正操作が必要なリスケジューリング問題として鉄道運転整理に着目し、運転整理案の立案に強化学習を適用した。鉄道運転整理に対して強化学習を適用し評価を行い、計画変更の立案に対する強化学習の有用性を示した。

今後の課題として、行動の選択肢が多い別のリスケジューリング問題に対する強化学習の評価検証が挙げられる。

参考文献

- [1] 崎川 修一郎, 佐藤 達広, 森田 豊久, 大田 健二, “運転整理知識を活用した制約プログラミングによる整理案作成方式”, 電気学会論文誌 C, Vol. 130, No. 2, pp. 332-342, 2010.
- [2] 電気学会・鉄道における運行計画・運行管理業務高度化に関する調査専門委員会, “鉄道ダイヤ回復の技術”, オーム社, 2010.
- [3] Open Data, “Timetables complete gfts”, <https://opendata.transport.nsw.gov.au/dataset/timetables-complete-gfts>, (2023-02-25 閲覧)
- [4] Matteo Hessel, et al., “Rainbow: Combining improvements in deep reinforcement learning” Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), pp. 3215-3222, 2018.