

## 強化学習環境における要求文書作成支援の評価

## Evaluating Assists in Specification Requirements for Reinforcement Learning Environment

森 渉<sup>†</sup> 原田 史子<sup>†</sup> 島川 博光<sup>†</sup>  
Mori Wataru Harada Fumiko Shimakawa Hiromitsu

## 1. はじめに

近年、自動車の自動運転など、状況が変化する非常環境を対象とした研究が増加している[1,2]. これらの課題解決には強化学習[3]が用いられることが多いが、環境作成には大きな労力を要する. 特に、非常環境のモデル化は難しく、経験の浅いユーザが制御対象の状態遷移を適切に記述するのは困難である.

本研究では、状態遷移図の要件文書作成を支援することで、強化学習の環境構築を容易にすることを旨とする. 具体的には、被験者に課題の要件文書を記述させ、与えるアドバイスの種類による影響を検証する. また、ChatGPTなどのLLMにも同じ情報を与え、生成された文書を人間の結果と比較する.

実験の結果、模範文書の提示や記述上のヒントが要件文書の品質向上に有効であることがわかった. また、LLMは経験の浅いユーザよりも良質な文書を作成できる可能性が示された.

## 2. 強化学習の環境作成のための状態遷移図作成

効率的な強化学習には、エージェントの行動や報酬を決定し、それをシミュレートする環境の構築が必要である. そのためには、良質な状態遷移図が不可欠となる.

状態遷移図は、制御対象の状態、遷移、行動を定義し、それらの関係を可視化するものである[4]. これにより、制御対象の複雑なふるまいを理解しやすくなり、シミュレーション環境の構築が容易になる. しかし、経験の浅いエンジニアが正確な状態遷移図を作成するのは難しい. そのため、まず制御対象の状態と遷移を自然言語で記述した要件文書を作成することが有効である. 良質な要件文書を基にすれば、適切な状態遷移図を作成し、強化学習の環境構築を支援できる.

## 3. 良質な状態遷移図作成の要件文書作成支援

## 3.1 テンプレートを用いた要件文書作成支援

強化学習環境の作成には良質な状態遷移図が不可欠だが、経験の浅いエンジニアには作成が難しい. 本研究では、要件文書作成支援による状態遷移図の質向上を目指す. 本手法の全体図を図 3.1 に示す. 本研究では、要件定義に不慣れたユーザを対象に、以下の2つの手法の効果を調査する.

1. 模範文書の提示：ユーザの書き方に近い例を示し、改善を促す.
2. 記述上のヒントの提供：要件文書の作成をサポートする.

本手法では、テンプレートを提供し、ユーザに要件文書を作成させる. その後、ユーザの書き方に似ている模範文書

<sup>†</sup>立命館大学院 情報理工学研究科

Graduate School of Information Science and Engineering

を提示し、ユーザに要件文書を変更させる. ユーザが作成した2つの要件文書を熟練者が精緻化し、変更前後の文書と精緻化した後の文書との編集距離を比較することで、支援手法の有効性を評価する. 変更後の文書で編集距離の変化が下がっていれば、本手法が有効であると判断できる.

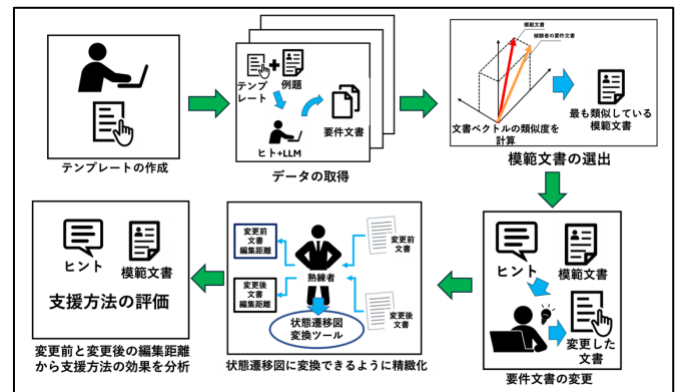


図 3.1

## 3.2 テンプレートの作成

本研究では、テンプレートを用いて良質な状態遷移図作成のための要件文書作成を支援する. テンプレートにより、不要な情報の記述や必要な情報の欠落を防ぎ、ユーザの設計を支援する. テンプレート作成のため、事前実験を実施した. 日本人男性大学生 12 人を対象に、パパ抜きの状態遷移図作成を目的とした要件文書を記述してもらった. 実験では、要件文書の書き方の例と例文を提示し、自由記述でパパ抜きの要件文書を作成させた. 作成された文書は熟練者が精緻化し、変更点を記録した. 共通する変更点を抽出し、良質な状態遷移図作成に適した特性とみなした. 初学者が良質な状態遷移図作成のための要件文書を書けるよう支援するため、本研究はこれらの特性をテンプレート化し、実験の説明書にも注意事項として書き加えた.

## 3.3 模範文書の提示

本手法では、初心者ユーザが記述した要件文書を Word2Vec で文書ベクトルに変換し、特定の次元空間に埋め込む. さらに、熟練者が正解例に変換した要件文書も同様に処理し、ユーザ文書との類似度を算出する. 最も類似度が高い正解例を模範文書として提示する. まず、Mecab で形態素解析を行い、単語に分割する. 次に、Word2Vec[5]の SkipGram を用いて単語を分散表現に変換し、文書内の全単語ベクトルの平均を文書ベクトルとする. さらに、コサイン類似度を用いて文書の類似度を算出する.

## 4 テンプレートを用いた要件文書

良質な要件文書作成支援手法の効果を調査するため、実験として、被験者に課題に対する要件文書を記述してもら

った。実験の目的は、テンプレートと注意事項を使用し、模範文書による変更が要件文書の書き方の理解度に与える影響を分析することである。実験では、まず被験者にテンプレートと注意事項なしの説明書を与え、ブラックジャックを題材にした要件文書を参考に、ババ抜きの状態遷移図の要件文書を記述してもらった。その後、テンプレートと注意事項ありの説明書を与えて同じ課題で要件文書を記述してもらった。

要件文書作成を支援する目的で、被験者に文書ベクトルの類似度がもっとも大きい模範文書を提示し、自らの要件文書を変更させる。変更前文書と変更後文書を精緻化し、編集距離を計算して有意差を検定した。編集距離に有意差があれば、模範文書を見せることが要件文書作成の支援に有効であることが示される。また、テンプレートと注意事項の有無についても同様に差分を算出し、有意差があればその有効性が示される。さらに、注意事項の順序についても検証し、有意差があれば重要な事柄は前に記述する方が理解度に貢献することが示唆される。

## 5. 支援手法の評価

### 5.1 模範文書提示による変更の効果

模範文書の提示が要件文書作成の支援に有効かを検証するため、変更前文書と変更後文書の編集距離を比較する。被験者が模範文書を見る前に記述した変更前文書と、模範文書を参考に修正した変更後文書を精緻化し、それぞれの編集距離を算出した。

図 5.1 に示すように、変更後の編集距離は変更前よりも小さく分布しており、有意水準を 0.05 としたウィルコクソンの符号付き順位検定の結果、 $p$  値は 0.001 となり、有意な差が確認された。編集距離の平均は変更前が 0.518、変更後が 0.404 であり、模範文書を提示することで要件文書の質が向上することが示された。この結果から、模範文書の提示は要件文書作成の支援に有効であることが分かった。さらに、LLM を用いた仮想被験者による文書を加えて

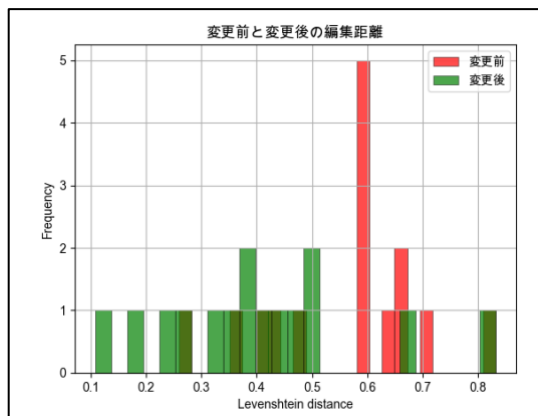


図 5.1

変更前と変更後の編集距離に差があるか検定した。有意水準を 0.05 とした検定の結果、 $p$  値が  $2.0 \times 10^{-4}$  となり、LLM が生成した文書を含んだ結果でも人のみで検定した場合と同様のことがいえる。LLM の結果を含めると平均編集距離が小さくなった。このことから、注意事項やテン

プレートの提供により、LLM は未習得のユーザよりも良い要件文書を書く可能性が示唆される。

### 5.2 注意事項の提示による要件文書作成の支援

テンプレートのみの説明書を与えた場合と、テンプレートと注意事項の両方を備えた説明書を与えた場合の要件文書を熟練者が精緻化し、それぞれの編集距離を比較した。

注意事項を提示した場合の平均編集距離は提示しないよりも小さかった。両文書の編集距離における分布の差を、検定を用いて統計的に確認した。有意水準を 0.05 とした結果、 $p$  値が  $9 \times 10^{-4}$  となり、有意な差が確認された。この結果から、記述上のヒントの提示は要件文書作成の支援に有効であるといえる。

さらに、人間は文書の冒頭を重要視する傾向があるため、説明書内のヒントを前に配置すると要件文書の理解が促進される可能性があるという仮説のもと、ヒントが説明書の初めにあるものを与えたグループ A と最後にあるものを与えたグループ B に分け、それぞれの変更前後の文書を精緻化し、精緻後の文書との編集距離の差を調べる。グループ A では変更前後の編集距離が 0.2 低下したのに対し、グループ B では 0.1 の低下にとどまった。これにより、変更前から変更後の文書における編集距離の変化に差がある可能性が示唆される。統計的な検証のためにマン・ホイットニーの U 検定を有意水準 0.05 で実施した結果、 $p$  値は 0.232 となり、両グループの編集距離の変化に有意な差は認められなかった。

## 6. おわりに

本論文では、強化学習の環境作成を容易にするため、良質な状態遷移図作成のための要件文書作成支援手法を提案した。提案手法では、テンプレートと複数の注意書きを用いてユーザに要件文書を作成させ、似た書き方の模範文書を提示し、文書を修正させる。変更前と変更後の文書、そして精緻化後の文書との編集距離を計算することで有効性を検証した。実験結果では、模範文書と注意書きが要件文書作成に有効であることが分かったが、注意書きの順序変更には有意差は見られなかった。LLM 生成文書を用いた場合、ヒトのみのデータよりも編集距離が小さくなる傾向が示された。今後、模範文書の書き方の特徴を説明することで、さらに効果的な文書支援が可能となることが期待される。

### 参考文献

- [1] Peide Cai, Hengli Wang, Huaiyang Huang, Yuxuan Liu, and Ming Liu. Vision-based autonomous car racing using deep imitative reinforcement learning, 2021.
- [2] Mehdi Masmoudi, Hamdi Friji, Hakim Ghazzai, and Yehia Massoud. A reinforcement learning framework for video frame-based autonomous car-following. IEEE Open Journal of Intelligent Transportation Systems, Vol. PP, pp. 1-1, 05 2021.
- [3] R.S. Sutton and A.G. Barto. Reinforcement learning: An introduction. IEEE Transactions on Neural Networks, Vol. 9, No. 5, pp. 1054-1054, 1998
- [4] Daniel R. Windle and L. Rene Abreo. Software Requirements Using the Unified Process: A Practical Approach. Prentice Hall PTR, USA, 2002.
- [5] Aurelien Geron. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, Inc., 2nd edition, 2019