

Pattern Explorer

パターンの記述に着目した探索ツールの開発

高梨大¹, 青木遥人²

概要

本論文では、分類と検索という2つの方法を通じて、過去に作成されたパターンを探索するツールを提案する。パターン・ランゲージはこれまで様々な領域で作成されてきた。本研究では、深層学習による文章埋め込みモデルを用いて、複数のパターンの記述をベクトル化することにより、パターンを記述内容に基づいて取り扱うことを可能にした。そして、記述内容に基づいてパターンを3次元空間に配置することで、より直感的な探索が可能なツール「Pattern Explorer」を制作した。ツールを制作した結果として、2つの活用方法と今後の可能性を見つけることができた。

1. はじめに

パターン・ランゲージは実践の支援ツールとして、これまで様々な領域で作成されてきた。その総数は数千を超え、その一つひとつがある領域における実践知を記述した情報である。そして、今後も様々な領域でパターン・ランゲージはつくられ続けるだろう。そうしたときに、膨大な数のパターンが存在することとなる。

大量の情報の中から目当ての情報にアクセスする際に、私たちは情報の探索を行う。そして探索を支援する方法として、分類と検索という2つの設計が存在する。例えば、図書館では大量の本をその内容に基づいて分類することで、利用者が目当ての本にアクセスできるようにしている。インターネットでは、Googleなどの検索ツールが存在することにより、特定のワードを打ち込むことで目当てのサイトにアクセスできるようになっている。

では、大量のパターンに対しては、私たちはどのように探索し、アクセスすることが可能なのだろうか。現時点では、パターンは特定の領域、例えばプレゼンテーション[1]や高齢者介護[2]などの、実践領域ごとに作成されている。そのため、領域ごとにパターンが分類されているため、実践したい領域さえ分かればアクセスが可能となる。検索による探索方法については、過去にソフトウェアパターンを検索するツールが作成されており、キーワードからパターンにアクセスすることができる[3]。

¹ 元慶應義塾大学 環境情報学部

² 慶應義塾大学 政策・メディア研究科

そこで、本論文では分類と検索の2つの方法を通じて、これまでに作成されたパターンを探索するツール「Pattern Explorer」を開発し、その活用方法や可能性を考察していく。

2. Pattern Explorer の設計方法

今回パターンを探索するツール、Pattern Explorer を開発するにあたり、記述そのものではなく、記述内容を取り扱えるようにした。これは、そのパターンの記述する実践がどのようなものなのかを踏まえた上での探索を目指すためである。実際に検索エンジンでも検索窓に特定のワードを入れたときに、そのテキストと合致するものではなく、意味に即した Web サイトが表示される。

そうした背景から、まずは深層学習による文章埋め込みモデルを用いて、パターンの Solution 部分の文章をベクトル化した。これにより、意味に基づいてパターンを取り扱い可能となった。文章のベクトル化には文章埋め込みモデルの一種である SimCSE [4]の日本語モデルである Japanese Simple-SimCSE [5]を用いた。SimCSE とは事前学習済み言語モデルと対照学習を用いた対照学習フレームワークであり、高品質な文ベクトルを生成することを可能とする。本ツールで用いたパターンのデータについて、本論文では、作成したツールによって新しい可能性を考察することを目的とした。そのため、井庭研究室から発表されている8種類のパターン・ランゲージを用いて、実験的にツールを動かした。

このようにして、意味内容に即してベクトル化されたパターンを用いて、次は Pattern Explorer の UI 部分の設計を行なった。UI をデザインする際には、利用者が検索を用いながら、直感的にパターンを探索することを目指し、2つの機能を実装した。1つ目の仕組みは、ベクトル化されたパターンを類似度に基づいて3次元空間上に配置する機能である(図1)。2つ目は、利用者が検索窓に入れたテキストの意味内容をリアルタイムにベクトル化し、それと近いパターンを5つ表示するという機能である(図2)。

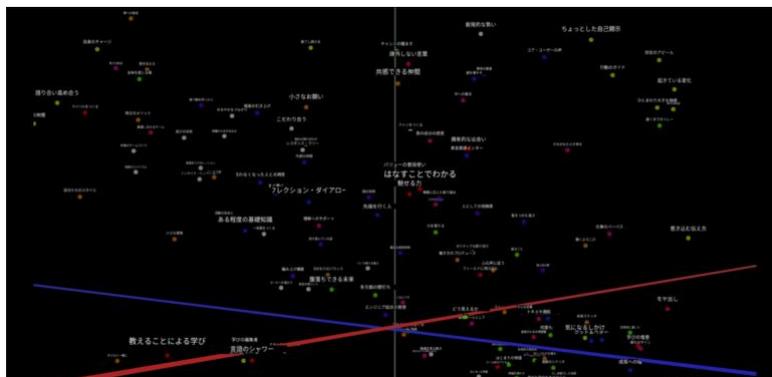


図1 ベクトル化され、3次元空間上に配置されたパターン

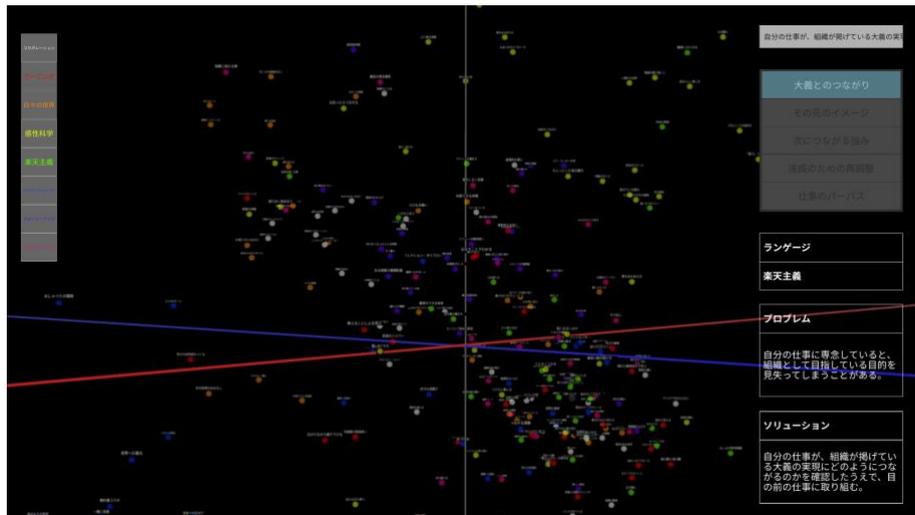


図2 検索窓にテキストを入れることによる探索の様子

以上の UI の実装には、Derivative 社のビジュアルプログラミングツールである TouchDesigner を用いた。バックグラウンドのシステムは Python を使用した。まず、1つ目の機能の実装には、先ほどの SimCSE によって 768 次元で表現されたベクトルを、PCA という次元削減手法を用いて次元削減を行い 3 次元データに変換した。そして、そのデータを、TouchDesigner を用いて 3 次元空間上に配置した。また、このように配置された空間内を、ユーザーがキーボード入力とマウス操作によって自由に移動可能にした。

2つ目の機能の実装は、ユーザーが検索窓に入力したテキストデータをリアルタイムでベクトル化し、近いベクトルを持つパターンを、コサイン類似度を用いて算出し、上位 5 つを表示するようにした。これにより、ユーザーはそのパターンを読み、その結果を踏まえて再度探索行えるようにした。

以上 2 つの機能を実装し、Pattern Explorer はローカル環境で動くパターンの探索ツールとして作成された。

3. Pattern Explorer の活用方法

本セクションでは、実際に Pattern Explorer に触れることで見えてきた活用方法を 2 つ示す。

まず、1つ目の活用方法は検索窓に自分の困っていることを打ち込むことによる悩み相談的な使い方である。これは、初めてパターン・ランゲージに触れる人が、無数に存在するパターン・ランゲージの中から自分の状況に適したパターンにアクセスする際に有効な使い方である。これは、今後パターンの数が増えるほどに、様々な状況に対応したパターンにアクセスできる可能性が上がっていくだろう。

2つ目の使い方は、パターンをつくっている人がこれまでに作成されたパターンを参考にするという活用方法である。これは、検索窓に現在記述しているパターンを打ち込むことによって、過去に作成された、似ているパターンにアクセスするという使い方である。これにより、現在作成しているものと同じ部分は参考にしつつ、異なる部分は意識的に書き分けることで、記述の支援ツールとすることができる。

以上2つは、検索という探索方法があることと、パターン・ランゲージを領域ではなく、パターンという単位で取り扱うことで生まれる活用方法だと考えられる。

4. おわりに

本論文では、パターンを記述の観点から探索を行うツールの開発と、その方法について説明した。また、ツールを実際に使うことで見えてきた2つの活用方法を示した。今回、実験的にツールを作成することで、パターンと深層学習を組み合わせることの可能性が見えてきた。

まず、パターンはこれまで、ある程度洗練された文章での記述が行われてきたため、整ったデータセットとしてすでに準備されている。そもそも、個々の分野での知識をランゲージ化しているため、近年発達している自然言語処理との組み合わせが可能となっている。

そして、パターンはランゲージを超えて、状況-問題-解決という共通の形式で記述がされているため、今後はその特徴を活かして、今回のツールの拡張や、新たなツール開発ができるかもしれない。

謝辞

まずは、本論文のシェパードをしていただいた島村恵理様に感謝します。論文に対して沢山のコメントをしていただき、ありがとうございました。そして、AsianPLoP2024で、この論文に対して多くのコメントを下さった WWSJ-3 の皆さんにも感謝します、ありがとうございました。

参考文献

- [1] 井庭崇・井庭研究室, 2013, 『プレゼンテーション・パターン: 創造を誘発する表現のヒント』, 慶應義塾大学出版会.
- [2] 金子智紀・井庭崇, 2022, 『ともに生きることば: 高齢者向けホームのケアと場づくりのヒント』, 丸善出版.
- [3] 中山弘之, 久保淳人, 鷲崎弘宜, 深澤良彰, 重要度算出に基づくソフトウェアパターン検索システム, コンピュータソフトウェア Vol.25, No.2, pp.114-134, 2008.
- [4] Tianyu Gao, Xingcheng Yao, Danqi Chen. SimCSE: Simple Contrastive Learning of Sentence Embeddings. EMNLP 2021 - 2021 Conference on Empirical Methods in Natural

Language Processing, Proceedings

[5] Hayato Tsukagoshi, Ryohei Sasano, Koichi Takeda. Japanese SimCSE Technical Report.