

共創支援を目的とした技術文書からの技術一価値表現の抽出

内田 貫太¹ 田口 亮¹

¹名古屋工業大学大学院 工学研究科 工学専攻 情報工学系プログラム
k. uchida.514@stn.nitech.ac.jp taguchi.ryo@nitech.ac.jp

概要

技術の構造化ツールとして VBridge というものがある。このツールは技術の理解や新しい価値の発想などに役立つことができるが、作成コストが高いという問題がある。そこで、技術文書などから VBridge を自動生成することが望まれる。本研究では、マルチタスク学習を適用させた汎用自然言語処理モデルを用いることで、論文から VBridge の要素を抽出することを目指した。また、その結果について考察を行い、検討すべき課題について整理する。

1 はじめに

近年、激しく変動する国際情勢において、企業には大きな需要の変化に対応できる力が求められている。そこで、大きな変化に対応するためには企業内外での部門間の垣根を越えた連携が必要不可欠となってくる。しかしながら、開発部門の人が社会的価値から着想を得るといことや、マーケティング部門の人が技術に対して深い理解を得るといことは難しい。この課題に対し、部門間の橋渡しとなる思考ツールとして VBridge (バリュー・ブリッジ) が提案されている[1]。VBridge は技術価値を視覚化したもので、下層の専門的な技術から、上層の社会的価値への道のりが順を追って表現される。これを用いることで、技術とその価値を共有することができ、部門間連携の助けとなる。しかし、VBridge を作成するためには、VBridge に対しての深い知識や、技術に対する理解が必要となり、誰でも容易に作成できるものではない。

本研究では、VBridge を作成するコストを削減するため、論文などの技術文書から VBridge を自動生成することを目指す。本稿ではその第一歩として、汎用自然言語処理モデルによる深層学習を用いて、VBridge の構成要素を抽出することを試みる。

2 VBridge

2.1 VBridge の概要

VBridge は技術が価値に結びつく過程を有向グラフとして表現したものである。下層から、なりたち (部品・機構)、ふるまい (自律機能)、はたらき (他律機能)、達成事項 (品質要素)、個々のよこび (要求項目)、みちすじ (共創プロセス)、ねがい (共創テーマ) の7階層に分かれており、各要素の因果関係が矢印で示されている。

[1]では、エアコンや冷蔵庫で用いられる帯電微粒子水の技術をヘアドライヤに応用した事例を用いて VBridge の利用方法を解説している。その VBridge を図1に示す。マーケティングや営業の担当者は、VBridge 上層にある社会的ニーズ (ねがい、みちすじ) や、「アレル物質から水素を抜き取る」や「ウイルスから水素を抜き取る」という、技術によって実現されている直接的な効果 (達成事項) は知っているても、その効果が得られるメカニズムや技術の構成要素については知らない場合がある。そこで、技術開発の担当者が、技術の構成要素 (なりたち) やその構成要素がどのように作用して現在の効果を生み出しているのか (ふるまい) を記述することにより、マーケティングや営業の担当者の理解を支援する。一方で、技術開発の担当者は、自社技術が具体

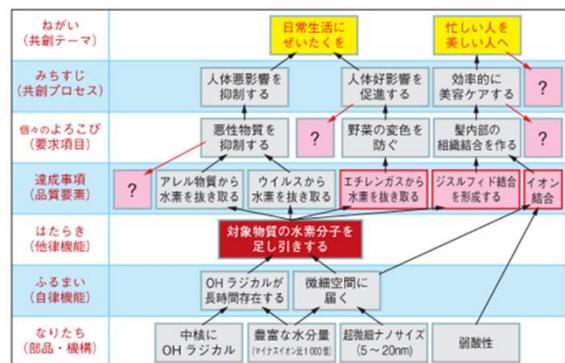


図1 VBridge の例 ([1]より転載)

的にどのような社会的ニーズを満たしているのかを知ることができるようになる。さらに、中間にある「はたらき」の層で、技術がもたらす効果を対象に依存しない表現で記述することによって、技術の可能性が広がり、空気清浄機で使われていた技術を、冷蔵庫やヘアドライヤに応用するという新たな発想が生まれやすくなるのではないかと論じている。

2.2 VBridge 自動生成手法の検討

VBridge の作成を自動化するためには、VBridge を構成する矢印や各階層をより具体的に定義する必要がある。本研究では、構成要素間の因果関係に着目し、あることがらを任意の層に当てはめるとすると、そのことがらを達成・解決したときに、達成できることをその1つ上層に位置付けるものとする。図1を例とすると、「対象物質の水素分子を足し引きする」ことが達成できれば、「アレル物質から水素を抜き取る」ことができるようになるかと理解できる。

このように技術の構造を解釈していくと、技術の全体像は「課題がある」、「課題を解決する」、「新しい課題が見える」のサイクルで論理が展開されていることが分かる。そうすることで、自ずと課題や解決法は、上段に行くほど広く一般的になっていき、下に行くほど狭く専門的になっていく。これはVBridgeの本来の目的である「技術理解の橋渡し」に相違ない。

ただし、技術文書から構成要素を抽出してVBridgeを作成するというアプローチでは問題が発生する可能性がある。例えば、ある技術文書の中で「異常データの収集は困難である」と「異常パターンを定義することは現実的ではない」という問題点が指摘され、その解決策として「正常データのみで学習する方法」が提案されているとする。この場合、一見して論理に齟齬はないと思われるが、実際はこの解決策でこれらの問題点が解決されるわけではなく、「異常データが必要不可欠である」というその技術文書では言及されていない別の問題点が解決されている。このように、技術文書には論理の飛躍が含まれる場合があるため、抽出手法のみでは完全な論理を表現するのは難しい。

従って、(1)技術文書からVBridge内に記述すべき構成要素を抽出し、(2)各構成要素間の因果関係を推定、(3)論理の飛躍を検出し補完することができれば、VBridgeの自動生成が可能になると考えられる。そこで本稿ではまず(1)の構成要素抽出を目標とする。

3 VBridge 構成要素の自動抽出

3.1 アプローチ

VBridge は容易に作成できるものではないため、データセットを大量に用意して学習させることでの精度向上は難しい。そこで、汎用自然言語モデルのファインチューニングというアプローチを採用する。また、本研究では最終的に2.2節で説明した(1)~(3)のタスクを実行することを想定している。深層学習では、各タスクを実行するモジュールを個別に最適化するよりも、マルチタスクとして同時最適化の方が、効率よく学習できる場合がある。

そこで本稿では、汎用言語モデルをマルチタスク学習によりファインチューニングすることにより、VBridge 構成要素の自動抽出を行う。

3.2 汎用自然言語処理モデル

汎用自然言語モデルとして有名なのはBERT[2]である。BERTは、Transformerというアーキテクチャを採用しており、このアーキテクチャは文章全体の依存関係を学習できるが、文章全体のトークン数が増加すると、計算量が $O(n^2)$ で増加してしまうという問題がある。つまり、技術文書のような長文に対して適用するのは難しい。そこで、長文の読解に適した言語モデルを使用することを考える。

Daiらはこの問題を解決するため、Transformer-XLというアーキテクチャを提案している[3]。このアーキテクチャは文章全体をセグメントに分割し、前のセグメントの情報を次のセグメントに引き継ぐことで長文の解析に対応している。

さらに、YangらはこのTransformer-XL[3]のアーキテクチャを使用したXLNetと呼ばれる汎用自然言語モデルを提案している[4]。このモデルはBERTではできなかった長文の解析ができることに加え、単語の順番をバラバラにして学習することでMasked LMにおける問題を解決している。

一方で、Liuらは単一タスクにおいて学習データが十分に得られない問題に対し、MT-DNN (Multi-Task Deep Neural Network) というモデルを提案している[5]。このモデルは、BERTに対して複数のデータセットでのマルチタスク学習を適用させることによって、BERT単体のスコアを上回っている。本研究では、ストックマーク株式会社が公開している日本語ビジネスニュース記事で事前学習を行った

表 1 構成要素の F1 スコア

手法	達成価値	将来価値	解決課題	将来課題	要素技術
シングルタスク学習	8.17	10.76	7.79	11.58	34.03
マルチタスク学習	9.26	10.79	6.67	8.33	35.00

XLNet[6]を使用する。また、[5]を参考に XLNet をマルチタスク学習によりファインチューニングする。

3.3 系列ラベリングによる構成要素抽出

本稿では、VBridge の構成要素抽出を、系列ラベリングタスクとして解く。各要素の説明を次に列挙する。

達成価値：その研究によって得られた価値

将来価値：その研究が将来的に達成したい目標

解決課題：その研究によって直接解決される課題

将来課題：将来的に解決したい課題

要素技術：その研究で用いられている要素技術

これらの要素は技術を説明する上で重要であり、VBridge を作成する上で欠かせないものとなる。しかしながら、「地球温暖化を解決する手法」といったように価値（地球温暖化を解決）、課題（地球温暖化）、要素技術（地球温暖化を解決する手法）が重複する場合があるため、各ラベルは重複可能とした。このような、重複した系列ラベリングを行う方法として、各ラベル独立に汎用言語モデルをファインチューニングする方法と、マルチタスク学習により汎用言語モデルをラベル間で共有する方法が考えられる。後述の実験では両者の性能を比較する。

4 実験

4.1 実験条件

インターネット上で公開されている情報処理学会第 84 回全国大会の講演概要[7]のうち、画像処理や深層学習に関する研究の概要 294 本を実験に用いる。前章で示した 5 つの要素をラベリングし、学習データを 250 本、評価データを 44 本としてデータセットを作成した。各ラベル独立にファインチューニングする場合と、マルチタスク学習によりファインチューニングする場合とを比較する。

4.2 実験結果と考察

表 1 が実験による各タスクの単体学習とマルチタ

スク学習の F1 スコアである。要素技術と達成価値についてはマルチタスク学習の方が抽出精度が高く、そのほかのラベルについては、シングルタスク学習の方が高精度となった。

今回の実験では、どのラベルにおいても良い結果は得られなかったが、この要因の一つに、予測結果に対する評価の手法が厳密すぎると考えられる。今回のデータセットにはアノテーションの揺らぎが生じている。例えば、「本研究では、地球温暖化を解決する手法を提案する。」といった文章に達成価値をラベリングすると考える。このとき、この研究の価値は「地球温暖化を解決できる」ことであるが、そのことを文章でラベリングする際には「地球温暖化を解決」、「地球温暖化を解決する」、「地球温暖化を解決する手法」などの複数の正解が存在してしまう。予測にも同じパターンが生じ、完全に一致しなければ評価が下がってしまうことになる。よって、ラベリングの意図を汲み取り、柔軟に対応できる評価手法が必要となる。

また、マルチタスク学習による精度の向上が見られなかった原因として、解決課題と将来課題のような似た傾向のあるラベリングタスクに対して、予測が難しくなったと考えている。

次に、実際の予測をもとにその傾向や課題について考える。表 2 が実際の予測と正解ラベルを一部抜粋したものである。各構成要素について以下のように考察する。

達成価値 「高速化」や「軽減」、「向上」など、ポジティブな単語を予測する傾向が強く見られた。これは価値を表す単語を予測できていると考えられるが、その主語となる単語までを予測することができていなかったため、文単位での予測が必要と考える。

将来価値 「密集」や「生産」、「コンピュータ」など、一般的な単語を予測する傾向が見られた。これは将来価値に含まれる単語の傾向とは一致するが、精度の向上のためには後の文脈から達成できていないことであることの推論が必要になると考える。

表2 各タスクの予測と正解ラベルの抜粋

要素	予測と正解ラベルの抜粋
達成価値	<p style="text-align: center;">predict</p> <p>…クラスタを計算環境とし、<u>流体の移流計算を効率的に高速化する手法を提案する。</u>…</p> <p style="text-align: center;">target</p>
将来価値	<p style="text-align: center;">predict</p> <p>…システムが<u>密集回避を促すことで感染リスクの軽減が期待される。</u>…</p> <p style="text-align: center;">target</p>
解決課題	<p style="text-align: center;">predict</p> <p>…<u>ロボットの音声認識精度が不十分で、スムーズな会話を継続することが難しい</u>…</p> <p style="text-align: center;">target</p>
将来課題	<p style="text-align: center;">predict</p> <p>…すべて記録する場合<u>多くの労力を要する。</u>…</p> <p style="text-align: center;">target</p>
要素技術	<p style="text-align: center;">predict</p> <p>…少数なデータを扱える<u>サポートベクター回帰を顔面領域皮膚温に適用し、</u>…</p> <p style="text-align: center;">target</p>

解決課題 他のラベルと比べてデータ数が少なく、出現しない文章も多く見られたため、予測が難しかったと考える。

将来課題 「要する」や「増大」、「かかる」など、ネガティブな単語を予測する傾向が強く見られた。これは達成価値と同様に、課題を表す単語を予測できていると考えられるが、その主語となる単語までを予測することができていなかったため、文単位での予測が必要と考える。

要素技術 「サポートベクター回帰」や「データベース」、「音声認識」など、専門的な単語を予測する傾向が見られた。これは要素技術に含まれる単語の傾向と一致しており、名詞のみでのラベリングが多かったため、他のラベルに比べてスコアが高かったと考えられる。

以上から、実験結果としてスコアは低かったものの、多くのラベルにおいて正解の傾向は見られており、評価手法やラベリング手法を改善することで精度が向上すると考えられる。具体的には以下の3つを今後の課題とする。

- ラベリングタスクでアノテーションの揺らぎに対応できる評価およびラベリング手法の改善

- より長い範囲で文を抽出する手法の検討
- 構成要素抽出以外の異なるタスクを用いたマルチタスク学習の実現

5 おわりに

本研究では、VBridgeの自動生成を目的として、その要素となるものを論文から抽出する手法を提案した。提案手法では、少ない自作データセットを活用するため、マルチタスク学習を適用させた汎用自然言語処理モデルを用いた。実験では、高い精度の結果は得られなかったものの、その予測の傾向から改善の余地があることを示した。

参考文献

- [1] 加藤雄一郎. 2016. 長期展望に立ったコア技術戦略, JTEKT ENGINEERING JOURNAL No. 1014.
- [2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. 2019. BERT: Pretraining of Deep Bidirectional Transformers for Language Understanding, Pro-ceedings of the

- 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume1 (Long and Short Papers), pp.4171-4186.
- [3] Zihang Dai, Zhilin Yang, Yiming Yang, Jaime Carbonell, Quoc Le, and Ruslan Salakhutdinov. 2019. Transformer-XL: Attentive Language Models beyond a Fixed-Length Context. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 2978-2988, Florence, Italy. Association for Computational Linguistics.
- [4] Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Ruslan Salakhutdinov, and Quoc V. Le. 2019. XLNet: Generalized Autoregressive Pretraining for Language Understanding. In NeurIPS 2019, pages 5754-5764.
- [5] Xiaodong Liu, Pengcheng He, Weizhu Chen, Jianfeng Gao. 2019. Multi-Task Deep Neural Networks for Natural Language Understanding, Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 4487-4496, Florence, Italy. Association for Computational Linguistics.
- [6] mkt3, 大規模日本語ビジネスニュースコーパスを学習した XLNet (MeCab+Sentencepiece 利用) モデルの紹介, <https://qiita.com/mkt3/items/4d0ae36f3f212aee8002>
- [7] 情報処理学会第 84 回全国大会講演論文集, https://www.ipsj.or.jp/event/taikai/84/ipsj_web2022/index.html