

## Word2Vec を用いた単語連想ゲームの試行実践と エンターテインメント分野への応用の検討

中矢 誠<sup>†</sup> 野口 克洋<sup>‡</sup> 富永 浩之<sup>†</sup>  
香川大学<sup>†</sup> 漫画家<sup>‡</sup>

### 1.はじめに

近年、機械学習による自然言語処理の発展が目覚ましい。特に、Word2Vec[1]を用いた単語の素性の獲得は、応用範囲の広い手法として注目されている。例えば、機械翻訳や、商品レビューの評価への応用発展が期待されている。

本研究では、この手法を、エンターテインメント分野に応用することを目指す。エンターテインメントコンテンツを構築するには、楽しめるように、ユーザの感性に寄り添う必要がある。そのために、Word2Vecによって獲得した単語の素性と人間の感性との差異を調べる必要がある。

### 2.単語連想ゲームによる試行実践

試行実践を行うために、単語連想ゲームをスマートフォンのアプリケーションとして実装した[2]。システム側が、キーワードを提示し、プレイヤーは、そのキーワードから連想した言葉を解答する。キーワードとプレイヤーの解答とのWord2Vecによる距離に応じて得点とする。1解答あたりに制限時間を設け、徐々に時間が短くなるように設定した。低い得点を出し続けた場合は、ゲームオーバーとなるようにすることで、プレイヤーに適当な解答を入力しないよう促した。

Word2Vecの辞書には、東北大学の乾・岡崎研究室が公開しているデータ[3]を用いた。このデータは、Wikipediaの情報から約100万語を学習している。アプリケーションは、iOSとAndroid向けにUnityで実装した。得点計算を行うサーバはRuby on RailsとPythonで実装した。

プレイヤーの連想に偏りができるだけ発生しないよう、工夫を行った。まず、画面上に表示する情報について、特定の単語の連想を誘発するキャラクター等のデザインを避けた。また、ルールに馴染めるよう、高い確率で「バナナ」をゲームの開始単語とするが、一定の割合で他の単語で開始されるように配慮した。ルール上、1

ゲーム内で同じ単語が二度以上は使えないようにすることで、少しでも多くの単語を収集するように設計した。

### 3.試行実践の状況と結果

2017年10月13日にiOS版を、翌14日にAndroid版をリリースした。約2ヶ月の間に、約100人によって310回プレイされ、14,310回の解答を得た。解答された単語の総数は、5,944であった。そのうち、辞書に含まれていない単語の総数は887(14.9%)であった。なお、開発者自身によるプレイデータが若干含まれているが、総数に比べると十分無視できる数である。また、同一プレイヤーによるプレイ回数については考慮されていない。そのため、単語の連想傾向に関して、若干の偏りが発生していることは否めない。

プレイヤーからのコメントもいくつか得たため、表1に示す。ポジティブな意見としては、評価が想定よりよくできているという意見があった。それに対して、ネガティブな意見としては、辞書の語彙力が足りない点や、特定の解答に対する得点に不満を覚えるものがあつた。

様々な単語から連想されている単語を、単語間の連想数順に上位30件だけ表2に示す。開始単語として頻出のバナナから連想されやすい単語のみに偏らず、動物や勉強など、小学校低学年の学習単語に近い分布の単語を中心に、広く分散して解答を得られた。下位になるにつれ、「アイドル」、「ダイエット」、「オリンピック」など、小学校中高年で使い始める単語が見られるようになった。さらに下位では、「犯罪」、「セックス」など、子供が日常的に使わない単語の割合が増えた。

### 4.辞書にない単語の分類

プレイヤーの解答のうち、約15%が辞書にない単語であり、ゲーム上では無効な入力として処理された。これについて、無効である原因別に分類したものを表3にまとめた。

表記ゆれは、「バイキンマン」と「ばいきんまん」といった、カタカナとひらがなの違いや、

Practices and Results of a Word Association Game  
as Entertainment Application using Word2Vec  
<sup>†</sup> Kagawa University  
<sup>‡</sup> Manga Artist

「尻」と「お尻」といった接頭辞の有無による違いなどによって、無効と判断されている。不正確な単語は、「電子マネー」あるいは「交通系電子マネー」とするべきところを、「交通マネー」と解答してしまう例が見られた。文章になり無効と判定された例としては、「頭がいい」や「海の神様」などがあった。このケースで、様々な単語からの連想があったものとして、「いい匂い」が挙げられる。それ自体が形容として用いられるために、プレイヤーが文章であるという自覚なく入力してしまうケースが多いのではないかと推察される。「青色申告」、「繁忙期」などは、辞書に収録されている語彙の不足により、無効と判定されていた。また、「けものフレンズ」や「妖怪ウォッチ」といった、近年のアニメ作品名などの新語も辞書に収録されていなかった。プレイヤーの中には、AI に対し、「は?」、「えー」、「クソ野郎」といった対話を試みたかのような解答を行っているケースがあった。これは、ユーザの解答に対する評価を、ゲーム画面上のキャラクターの表情やセリフでビジュアル化していたために誘発された解答だと考えられる。

## 5.考察

一部の単語については、評価に不満を感じる声があった。これは、Word2Vec 上での単語の素性が、同じカテゴリのものほど関連度が高いとされやすい傾向があるためだと考えられる。例えば、「バナナ」と「トマト」や、「黄色」と「赤」は、それぞれ高い関連度とされるのに比べ、「バナナ」と「黄色」や、「トマト」と「赤」では、関連度が低くなる。また、Wikipedia の記事から学習した辞書を用いたため、文章上に表顕しにくい色や季節感、五感情報などは、プレイヤーの感性と顕著な差異が見られがちである。例えば、「海」と「魚」に比べ、「海」と「夏」では、関連度が低くなる。

辞書の語彙については、不満に思う声と、特に問題ないという声に分かれた。日常的に使う語彙が人により異なるため、それが専門的な語群であっても、自分の興味のある領域の単語が評価されないことに不満を感じるパターンが少なくないようである。

## 6.おわりに

スマートフォンのアプリケーションとして、単語の連想ゲームを実装してリリースし、プレイヤーの利用データを集め、分析した。今後の課題として、さらにプレイヤーの利用データを蓄積し、統計的に関連があることが明らかな単語の評価を改めるなど、データを活かし、プレイヤー

の感性に寄添う仕組みづくりが必要である。その上で、データの汚染、間違った学習がなされる可能性などを考慮し、表3のような原因ごとの分類と個々の解決策の模索を継続的に続ける必要がある。

こうしたデータの分析と蓄積によって、課題を解決していくことで、単語間の関連性を利用した他のゲームへの応用が期待できる。具体的には、児童向けの教育ゲームや、クロスワードパズル、ワードウルフなどが挙げられる。

表1 プレイヤからのコメント

- ・せっかくひねった解答をしたのに評価が低かった
- ・思ったよりちゃんと単語の関連度を評価しているように感じる
- ・固有名詞に弱い
- ・知らない単語が多すぎる。白物家電とか中東戦争ですらダメだった。浅い連想だけが評価される。
- ・長く続けていると単調で辛い

表2 連想数の多い単語

単語	連想数	単語	連想数	単語	連想数
甘い	40	学校	25	映画	21
車	40	白い	25	水	20
夏	34	お金	25	長い	20
海	32	バナナ	23	太陽	19
緑	30	果物	23	仕事	18
赤	28	ケーキ	22	楽しい	18
動物	28	冬	22	サッカー	18
勉強	27	りんご	22	いちご	18
野菜	27	山	21	かわいい	18
魚	26	花	21	美味しい	18

表3 無効な解答の原因

無効な原因	内容
表記のゆれ	・ひらがな/カタカナ/漢字の違い ・全角英数/半角英数の違い ・接頭辞/接尾辞の有無
不正確な単語	・正しくない表記
誤字	・打ち間違い ・変換ミス
文章	・接続詞の混入 ・修飾語の有無
語彙の不足	・新語 ・固有名詞 ・その他
AI との対話	・口語

## 参考文献

- 1) Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean: Efficient estimation of word representations in vector space, CoRR, abs/1301.3781, (2013).
- 2) Aqutras Inc.: 連想力 UP! ロジカルバナナ, <https://itunes.apple.com/jp/app/id1291769646?l=ja>.
- 3) 鈴木正敏: 日本語 Wikipedia エンティティベクトル, [http://www.cl.ecei.tohoku.ac.jp/~m-suzuki/jawiki\\_vector/](http://www.cl.ecei.tohoku.ac.jp/~m-suzuki/jawiki_vector/).