

AIによる料理創作支援システム

薦田 一慶[†] 岡 誠[†] 兼子 毅[†] 森 博彦[†]

東京都市大学大学院総合理工学研究科情報専攻システム情報工学領域[†]
東京都市大学情報工学部知能情報工学科[‡]

1 背景

近年、発想を支援する研究は数多く行われている。最近ではコンピューターを用いてデザインの支援や絵を考える際の支援、曲のコードやフレーズを提案し、発想や創作の支援をしてくれるものがある。

また近年、技術の発展により人工知能などのコンピューターが身近な存在となり、人間と対話をとる機会が増えた。更にアドバンスチェスと呼ばれる人とコンピューターが組んで戦うチェスの競技ではチェスの王者や高性能のコンピューターより一般人とコンピューターが組んだ方が強いという実績があり、人とコンピューターの協調には可能性がある。

上記のような支援を行う研究は多くされているが協調をしながら作業を支援する研究は少ない。

2 関連研究

2.1 人の発想支援について

IBM[1]は自社の人工知能である Watson を用いてシェフ・ワトソンを開発した。これはどんな食材をメインとするか、調理方法、現在の気分の三つのキーワードを入力する。それを元に使用者のイメージする料理にもっとも近い食材の組み合わせからタイトルを考案し、調理方法を提案するものである。これはユーザーが最初以外関わる要素がなく途中でユーザーが意見を挟むことはできない。

2.2 単一モデルでの複数言語間翻訳について

Melvin ら[2]は複数言語間での翻訳を行うために入力文の先頭に翻訳先の言語を指定するトークンを挿入することで単一モデルでの複数言語間での翻訳に成功している。また学習の際に直接的に関係のなかった言語ペア間でも翻訳に成功しており、構造が似ている言語間では一定の品質があるが、あまり似ていない言語間での品質は下がった。

2.3 文脈文を考慮した対話における機械翻訳について

Rachel ら[3]は文脈文を考慮した対話における機械翻訳をマルチ Encoder モデルと文脈文と入力文を繋げて入力するシングル Encoder モデルの比較を行った。テストデータは四つジャンルの映画の対話を元に作成されたテストセットを用いた。マルチ Encoder を用いた場合に一番良い結果を示した。

Collaboration with AI in cooking

[†]Komoda Kazuyoshi, System Information Engineering,
Tokyo city University Graduate School

[‡]Makoto Oka, Hirohiko Mori, Takeshi Kaneko, Department of
Intelligent Systems, Tokyo City University

3 研究目的

関連研究では AI との対話や協調作業での支援を行うもの、人工知能を用いて創作の際の提案をしているものはあるが何かの作成途中で現在の情報を考慮した提案、支援をするものは数少ない。

本研究では人間とコンピューターの対話を用いて人間が最も関わる人が多いであろう創作活動である料理の創作支援を目標とする。以下の事を利用の想定している。

こんな材料があるのだけれど何を作ろうか、料理を行っている最中に何か別の物を作りたい、アレンジをしたいといったレベルの違う要望に対する助言を提示することを想定している。このような使用段階でレベルの変わる質問に対して現在までの質問や情報を利用することで現状から考えられる答えを出力とする。また文脈情報を利用し、レベルの変わる質問に対する出力を行うには質問ごとに学習器を用意するのでは文脈情報を取り入れることが出来ないため、一つの学習器で学習を行う必要がある。これを元に料理の創作を行うことを想定している。

本研究ではその創作支援を行うための前段階として一つの学習器で前の質問の情報を文脈情報として保持して料理に関しての複数レベルでの質問に対する答えを与える事で対話を行うことを目的としたシステムの設計、提案を行う。

4 学習用データについて

本モデルを動かす目的は料理という場面での支援であるため学習に使用するデータは料理に関するものである。

想定している質問の種類としては以下の表 1 のような想定をしている。

表 1 質問の種類と例

質問	質問に対する答え
～があるのだけれど何が作れる？ (作るものへの質問)	料理のタイトル
～ってどんな料理？ (どんな料理かの質問)	料理についての説明
～をアレンジしたい、何かいいものはある？ (何か材料を追加するための質問)	その料理によく使われている材料のアドバイス

このような質問を想定しているため、学習用データはタイトル、材料、料理の説明のデータを使用し

て質問（入力データ）、質問の答え（教師データ）を作成する。また本研究では対話をしながらシステムの使用をすることを想定しているため、文脈情報を含んだ学習用データが必要である。この文脈情報とは一つ前の質問の出力とその料理で使用された食材とする。そのための学習用データは上記の質問、質問の答えを使用し、通常の Attention seq2seq を用いて作成する。データは Cookpad [4]より提供されているレシピを元にクリーニングを行い、学習に適した形にする形態素解析、パディングを行う。

5 提案システム

5.1 単一学習器での質問の分類

今回想定している質問は大きく分けて、作るものへの質問・どんな料理かの質問・何か材料を作るための質問の計三種類である。

これらを一つの学習器で分類しつつ、質問に対応するために Melvin ら[2]の手法を用いて入力先頭に質問の種類を指定するトークンを図2の問のように挿入する。

5.2 提案モデル

提案システムの想定としては質問をそのまま入れることを想定しているため、時系列データとしての処理を行うべきである。下記のように通常の Encoder, Decoder を元に入力データを文脈文と質問である入力分を繋げて入力する事で文脈文に対応させモデルである。

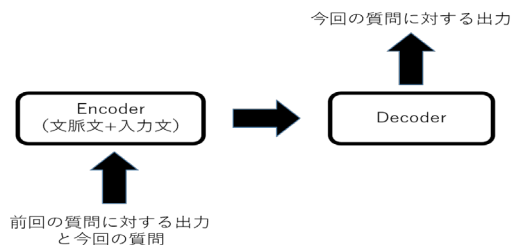


図1 提案モデルの図

提案モデルと 5.1 の手法を組み合わせる事で複数種類の質問に対する出力を文脈情報を保持しながら一つの学習器で得ることができると思われる。

6 学習結果

6.1 文脈情報無しでの学習結果

問	<in>ハンバーガーを何かプラスでアレンジしたいのだけどいい材料ある?>>>
教師	マヨネーズチーズアボカド>>>
出力	トマトビーマンマヨネーズ>>>
問	<ex>明太子焼きそばってどんなもの?>>>
教師	和風の明太子焼きそば>>>
出力	和風の明太子焼きそば>>>
問	<ti>醤油ごま油片栗粉があるのだけど何が作れる?>>>
教師	かんたん油淋鶏>>>
出力	かんたん油淋鶏>>>

図2 トークン挿入時の学習結果の一部

実際に学習を行った結果以下の図2のような出力が得られたため、質問の種類を指定するトークンを入

れることは効果的であると考えられる。

6.2 文脈情報ありでの学習結果

提案モデルでは質問と文脈文を入力データとし、質問の答えを教師データとして学習を行った。提案モデルの学習結果は以下の図3である。

問	<ti>クリームチーズラムレーズン卵があるのだけど何が作れる?>>>
教師	抹茶チーズケーキ>>>
出力	梨とクリームチーズのカプレーゼ>>>
問	<ex>酢豚って何?>>>
教師	本格的な中華酢豚の味です!>>>
出力	野菜がたくさん入った酢豚です。>>>
問	<in>塩焼きそばを変化させたいけどいいものない?>>>
教師	ごま油 みりん 醤油>>>
出力	ナンプラー ごま油>>>

図3 提案モデルの学習結果の一部

提案モデルでは文脈文を取り入れることで、図3では一部でしか紹介できていないが取り入れない場合よりも状況を考慮した出力を得ることが出来た。具体的にはアレンジの際に使用した食材が出にくくなり、類似した料理からのアドバイスが得られた等がある。しかし料理の説明(<ex>の部分)では、情報量が増えたためか、共通した食材を用いたほかの料理の説明がなされるといった問題点も見つかり、この点についてはデータの工夫やモデルの改良を行う必要がある。

7 今後の展望

提案システムの入出力では定量的な評価が行えないため、人の目による評価実験を行う必要がある。また評価実験で十分な評価が得られ次第、実際に人とシステムで協調をし、料理の創作支援が行えるかの実験を行い、システムの有効性の評価を行う。

参考文献

- [1] <https://www.ibm.com/watson/jp-ja/>
- [2] Melvin Johnson, Mike Schuster, Quoc V. Le, Maxim Krikun, Yonghui Wu, Zhifeng Chen, Nikhil Thorat, Fernanda Viégas, Martin Wattenberg, Greg Corrado, Macduff Hughes, Jeffrey Dean “Google’s Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation” Transactions of the Association for Computational Linguistics, vol. 5, pp. 339–351, 2017.
- [3] Rachel Bawden, Rico Sennrich, Alexandra Birch, Barry Haddow “Evaluating Discourse Phenomena in Neural Machine Translation” Proceedings of NAACL-HLT 2018, pages 1304-1313
- [4] <https://cookpad.com/>