

ユーザのライフログに対する健康アドバイスの自動生成

粟村 誉[†] 岡 照晃[†] 荒牧 英治[‡] 河原 大輔[†] 黒橋 禎夫[†]

[†] 京都大学大学院 情報学研究科

[‡] 奈良先端科学技術大学院大学 情報科学研究科

awa@nlp.ist.i.kyoto-u.ac.jp, oka-teruaki@i.kyoto-u.ac.jp,

aramaki@is.naist.jp, {dk, kuro}@i.kyoto-u.ac.jp

1 はじめに

日本に代表される先進諸国での高齢化が問題視されている。厚生労働省の予測によると、日本の総人口における高齢者の割合は2050年までに40%に上昇し¹、国民が十分な医療リソースを享受できなくなる恐れがある。このため、今後は人々が自らの健康管理をある程度、能動的に行い、自らの健康を維持していくことが求められる。しかし、健康管理を能動的に行うことは難しく、専門家のサポートが求められることも多い。

そこで、我々は、自然言語処理を用いた健康アドバイスの自動生成を行うシステムの開発を行っている。提案システムでは、ユーザが執筆したライフログ（その日行なった食事や運動についての日記）に対して、その人の状態に対応したアドバイスを生成することを目指す。

文生成や対話に関する研究は昔から盛んに研究されている分野であり [1]、テンプレートに基づく生成方式 [2] やルールに基づく生成方式 [3]、統計的な生成方式 [4] などがある。これまでも健康に関するシステムとして、定型文を用いた健康管理アプリケーション² や、対話形式などでメンタルヘルスケアを行うシステム [5] が報告されている。

本研究では、まずユーザのライフログとそれに対する専門家のアドバイスを収集した（ライフログ・アドバイスコーパス）。アドバイスの生成手法として、入力されたライフログとコーパス上のライフログとの Bag-of-words の類似度の高いアドバイス文を選択する。選択したアドバイス文を入力ライフログに対するアドバイスとする。本研究では、各アドバイス文の内容を、食事に関してそれぞれ概念としては何を伝えようとしているか（例えば「野菜を食べる」「3食とる」など）

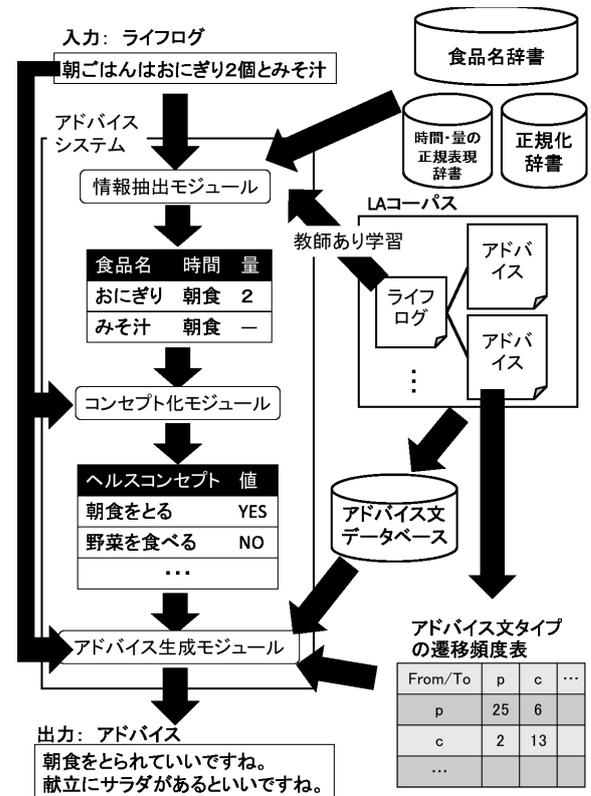


図 1: アドバイス生成の概要

という観点から分類した。この概念を本稿ではヘルスコンセプトと呼び、アドバイス文選択の際に類似度と同時に用いるアドバイス生成手法を提案する。

我々のアドバイス生成システムの全体の概要図を図1に示す。アドバイス生成システムの概要については文献 [6] を参照してほしい。本研究では図1のうち、アドバイス生成モジュールについてのアドバイス文選択に関する提案手法を示し、実際の生成アドバイスを比較する。

¹<http://www.mhlw.go.jp/wp/hakusyo/kousei/14/>

²あすけん：<http://www.asken.jp/>

2 ライフログ・アドバイスコーパス

本研究では、ユーザが執筆したライフログから必要な情報を抽出し、適切な健康アドバイスを自動生成することを目的とする。このようなシステムの開発・分析に必要なリソースとしてまず、男性 10 人、女性 10 人にそれぞれ連続した約 20 日間の健康に関するライフログの執筆を依頼した。これにより合計 400 件のライフログが集まった。また、ライフログに対するアドバイスのデータベース作成と評価のため、管理栄養士 2 名にアドバイスを付与しやすいライフログへの健康アドバイスの執筆を依頼した。その結果、400 件のライフログのうち、138 件に対して各 2 人分のアドバイスが付与された。さらに、この 138 件のライフログと 276 件のアドバイスについて、アドバイス文が各ライフログの何行目に関して言及しているかを各アドバイス文に対して人手で付与した。これをアドバイスリンクと呼ぶ。本研究では、この 138 件のライフログ・アドバイスをライフログ・アドバイスコーパスと呼び、利用する。

3 ヘルスコンセプト

人間が実際に行うアドバイスにはほぼ無限のバリエーションがあると思われるが、伝達したい意味レベルでの内容（例えば、「野菜を食べる」）を考えると有限であると考えられる。この意味レベルでの抽象的な概念を本稿ではヘルスコンセプトと呼ぶ。例えば、「献立にも野菜をたっぷりとり入れてみてくださいね。」というアドバイス文には「野菜を食べる」というヘルスコンセプトがあるとみなす。このようにしてライフログ・アドバイスコーパスのアドバイスを分類した結果、以下の 17 種の食事に関するヘルスコンセプトが得られた。

- 健康に気を使う
- 食生活がよい
- 3食とる
- 朝食をとる
- 食事献立がよい
- 体を温める
- 量を控える
- 油ものを控える
- 甘いものを控える
- 炭水化物を控える

ヘルスコンセプト	値	ライフログ
朝食をとる	NO	今日は忙しくて朝は抜き。お昼にリンゴを食べた。最近野菜をあまり食べていない気がするので気を付けようと思います。
果物を食べる	YES	
野菜を食べる	NO	
ヘルスコンセプト	値	アドバイス
朝食をとる	NO	朝ごはんを食べないと、体は脂肪を蓄積しようとう働きます。献立にも野菜をたっぷりとり入れてみてくださいね。
野菜を食べる	NO	

表 1: ヘルスコンセプトアノテーション例

- 生姜をとる
- 汁物を飲む
- ビタミンをとる
- タンパク質をとる
- 果物を食べる
- 野菜を食べる
- 魚を食べる

さらに、それぞれのヘルスコンセプトは、そのコンセプト名となっている行動を実際に行ったかどうかという情報（事実性情報）を持ちうる。この事実性として以下の 3 つを扱う。

- YES：実際に行った
- NO：実際に行わなかった
- Unknown：言及されていない

ヘルスコンセプトとその値の具体的な付与例を表 1 に示す。

本研究では、ライフログ・アドバイスコーパスを構築する際に全てのアドバイス文にヘルスコンセプトを手で付与しておき、入力されたライフログに対してヘルスコンセプトを自動付与することで、そのヘルスコンセプトに対応するアドバイス文集合を抽出できる。これにより、ライフログに対応するアドバイス文集合が得られ、その文集合からアドバイス文を選択することによってアドバイスの生成を実現する。

4 アドバイス文選択手法

3章で定義したヘルスコンセプトを用いてアドバイスを生成することを考える。ある入力ライフログのヘルスコンセプトとその値がわかった時、そのヘルスコンセプトに紐付いているアドバイス文集合からアドバイス文を選択することで、ライフログの抽象的な情報に則したアドバイスを生成できると考えられる。本研究では、候補となるアドバイス文集合からのアドバイス文の選択について次の 4 つの手法を比較する。それぞれの手法の概要を図 2 に示す。

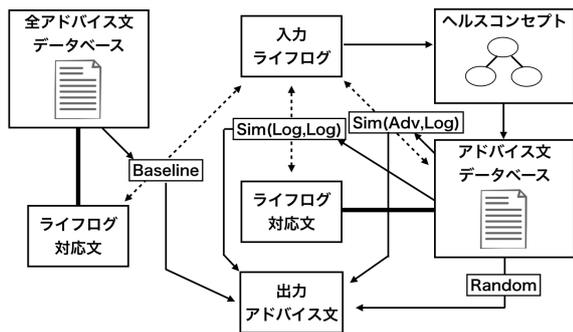


図 2: アドバイス文選択手法．太線はアドバイスリンク，実線はデータと処理の流れ，破線は類似度の計算を表す

- Baseline

ヘルスコンセプトを使用しない手法であり，本研究のベースラインである．入力ライフログが与えられたとき，ライフログ・アドバイスコーパス中のすべてのアドバイス文に関して，2章で作成したアドバイスリンクによりリンクしたライフログ文と入力のライフログとの類似度を計算し，類似度の高い順にアドバイス文を選択する．

- Random

入力されたライフログに自動付与されたヘルスコンセプトを使用する．ランダムに1つヘルスコンセプトを選択し，選択したヘルスコンセプトに紐づくアドバイス文集合の中からさらにランダムに1文選択する．

- Sim(Adv, Log)

入力されたライフログに自動付与されたヘルスコンセプトを使用する．付与されたヘルスコンセプトに紐づくアドバイス文集合の中で，アドバイス文と入力ライフログとの類似度が最も高いものを選択する．

- Sim(Log, Log)

入力されたライフログに自動付与したヘルスコンセプトを使用する．付与されたヘルスコンセプトに紐づくアドバイス文集合の中で，アドバイス文にリンクしているライフログ文と入力ライフログとの類似度が最も高いものを選択する．

類似度には Bag-of-words のコサイン類似度を使用した．Sim(Log, Log) での生成をヘルスコンセプトを考慮せず行う手法が Baseline となる．

各手法はそれぞれ適切だと判断したアドバイス文を1文ずつ選択していくことによってアドバイス文を生成

する．すべての手法に関して，アドバイスの合計文字数が100字以上となった時点でアドバイスの出力を終了する．選択できるアドバイス文がなくなった場合は100字以上でなくても終了する．これは，実際のアドバイスの文字数が約100字であることと，本研究で用いた自動評価尺度がシステムの出力の長さに依存するため，その影響を抑えるためである．

5 アドバイスの生成実験

5.1 実験設定とデータ

3章で定義したヘルスコンセプトを用いて，ライフログに対するアドバイスの生成実験を行った．データにはライフログ・アドバイスコーパスを使用した．入力ライフログのヘルスコンセプトは前述のシステムによって自動付与した．アドバイス文の選択については4章で述べた手法を比較する．

出力アドバイスの評価は，自動要約タスクで主に用いられる ROUGE-1[7] を使用した．正解データは，各ライフログに実際に付与されている2人分のアドバイスを用いた．ROUGE-1 の計算は単語単位で行う．そのため，正解のアドバイスおよび生成されたアドバイスを単語分割する必要がある．単語分割には，日本語形態素解析器 JUMAN³ を使用した．実験は5分割交差検定で行なった．ランダムな要素のある手法を含むため5回の試行の評価の平均を最終的なスコアとした．

また，図1で示したアドバイス文タイプの遷移頻度表を用いたアドバイス文の順序の並べ替えは行わない．これは，評価に単語の1-gramのみを使用する ROUGE-1 を使用しており，アドバイス文の順序が評価に影響しないためである．

5.2 実験結果と考察

アドバイスの生成実験の結果を表2に示す．表2より Sim(Log, Log) でのアドバイス生成で他の手法に比べ評価値の向上が見られた．Sim(Adv, Log) で評価値の向上はあまり見られなかったが，これは人間が必ずしもオウム返しのようにライフログ中の記述と同じようなアドバイスをするわけではないためと考えられる．また，Baseline と各手法との比較から，ヘルスコンセプトを用いてアドバイスを生成することによってアドバイスの評価値が向上することを示せた．

³<http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

Baseline	Random	Sim(Adv, Log)	Sim(Log, Log)
0.222	0.247	0.258	0.267

表 2: アドバイスの評価

Baseline ではヘルスコンセプトの考慮をしていないため、ライフログの内容と全く関係のないアドバイスが生成されることがある。例えば、野菜を食べている記述のあるライフログに対して「献立にも野菜をたっぷりとり入れてみてくださいね。」といった野菜を食べていないライフログについてのアドバイス文が選択されることがある。また、Random ではアドバイス文を選択する際にライフログの詳細な内容を考慮しないため、ライフログで言及のない食事（味噌汁、パスタなど）や行動に関するアドバイスが生成されてしまう。Sim(Log, Log) ではこれらの問題が改善され、ライフログの内容にあった適切なアドバイスが生成されていたことがわかった。

本研究では生成したアドバイスの評価として、自動評価尺度である ROUGE-1 を使用した。この評価尺度は、正解のテキストの単語がシステムの出力にどの程度含まれるかというものである。本研究では図 1 のアドバイス文タイプの遷移頻度表を用いたアドバイス文の順序の考慮は行わなかった。しかし、生成されたアドバイスの中にはそれぞれのアドバイス文が適切であっても、食事に関するアドバイス文が食事の時系列順に並んでいないなど、全体を通して不自然に感じるものもある。このようなアドバイスに関しても、ROUGE-1 の評価では単語のみを参照するため特に評価に影響がないという問題がある。そのため、文の順序を考慮した自動評価や人手での評価などを検討する必要がある。

また、本研究ではある 1 ライフログを入力として使用したが、実際のサービスとして考えるとある 1 日だけを対象としてアドバイスを生成するのは適切でない。例えば、毎日油ものを控えている人がある日油ものを摂取した場合、頭ごなしに批判されてしまうと普段の行いを理解してもらえてないものだと感じてしまう。逆に野菜を毎日食べていたり、油ものを控えていたりする人に対しては、毎日継続していることについて理解をしたアドバイスをすべきであると言える。これらのアドバイスの生成のためには、入力ライフログだけでなく n 日前までのライフログを参照する必要がある。

6 おわりに

本研究では、ユーザの執筆した健康に関するライフログについて適切な健康アドバイスを生成することを目的とし、システムの開発を行った。その際に、入力ライフログに関する抽象的な概念とコーパス上のライフログ・アドバイスとの類似度を用いてアドバイス文を選択する手法を提案した。結果として、提案手法において生成したアドバイスの評価値の向上が見られ、特に抽象的な概念を考慮し、入力ライフログと類似するライフログにリンクしたアドバイス文を選択する手法において、ROUGE-1 でベースラインから 0.045 ポイントの向上が見られた。今後の課題として、過去 n 日間のライフログを対象とした長期間の行動に対するアドバイスの生成にも取り組みたい。

謝辞

本研究は、革新的イノベーション創出プログラム (COI STREAM) 「活力ある生涯のための Last5X イノベーション」の支援を受けた。

参考文献

- [1] Ehud Reiter and Robert Dale. *Building Natural Language Generation Systems*. Cambridge University Press, New York, NY, USA, 2000.
- [2] Stephanie Seneff. Response planning and generation in the mercury flight reservation system. *Computer Speech & Language*, Vol. 16, No. 3-4, pp. 283-312, October 2002.
- [3] Fabrizio Morbini, Eric Forbell, David DeVault, Kenji Sagae, David Traum, and Albert Rizzo. A mixed-initiative conversational dialogue system for healthcare. In *SIGDIAL*, pp. 137-139, Seoul, South Korea, July 2012.
- [4] Alice H Oh and Alexander I Rudnicky. Stochastic natural language generation for spoken dialog systems. *Computer Speech & Language*, Vol. 16, No. 3-4, pp. 387-407, 2002.
- [5] Fabrizio Morbini, David DeVault, Kenji Sagae, Jillian Gerten, Angela Nazarian, and David Traum. FLoReS: A forward looking, reward seeking, dialogue manager. In *Proc. 4th International Workshop on Spoken Dialog Systems*, pp. 151-162, Paris, France, November 2012.
- [6] 岡照晃, 栗村誉, 荒牧英治, 河原大輔, 黒橋禎夫. おしゃべりけんこうノート: 管理栄養士・インストラクターのアドバイスに基づく健康アドバイスシステム. 言語処理学会第 22 回年次大会, March 2016.
- [7] Chin-Yew Lin. Rouge: A package for automatic evaluation of summaries. In *Proc. ACL-04 Workshop on Text Summarization Branches Out*, pp. 74-81, Barcelona, Spain, July 2004.