

# 意図推定法を用いた協調型強化学習に基づく推薦システムの提案

椿本 樹矢

指導教員：小林 邦和

## 1 はじめに

マルチエージェント環境下における協調行動の獲得に関して強化学習 [1, 2] を用いた意図推定法 [3] が著者らによって提案されている。この手法は汎用的なものであり、マルチエージェント環境以外の問題に対しても適用が可能である。本論文では推薦システムを取り上げる。

推薦システムは、利用者にとって有用であると思われるアイテム (情報、商品) などを提示するシステムである。最も基本的な手法として、内容ベースフィルタリングと協調フィルタリングが挙げられる [4, 5, 6]。また、特定の推薦問題に対して、様々な推薦システムが提案されている [7, 8, 9]。これらの手法はその構成上事前に嗜好データが必要、もしくはある問題に特化した手法であり、他の新たな問題への適用が困難である。

本論文では、意図推定法を推薦システムへ適用することで意図推定法の汎用性を検証する。特にここでは、事前データの少ない問題を含む、小規模な推薦問題を対象とする。提案システムは意図推定法の汎用性を利用し、様々な推薦問題に対し適用可能である。また、提案システムはユーザーの嗜好を学習することで推薦を行う度にユーザーに合ったシステムへと変化する。実用例として献立推薦システムを取り上げ、計算機シミュレーションを通して有用性の検証を行う。

## 2 提案システム

本システムは、特徴抽出部、推薦学習部、嗜好学習部の3つで構成される。システム構成を図1に示す。

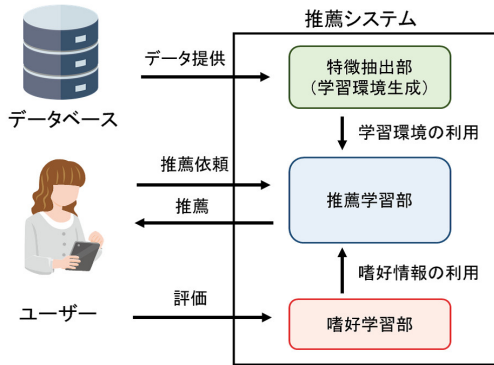


図1 提案システムの構成

### 2.1 特徴抽出部構成

特徴抽出部は、アイテムの特徴を自己組織化マップを用いて抽出する部分である。自己組織化マップは教師なし学習の手法であり、データの持つ特徴ベクトルを基にデータを分類する手法である [10]。データベースから与えられる多次元のデータを持ったアイテムデータを二次元のグリッド空間上に配置する。グリッド空間上に配置されたアイテムをゴール、二次元グリッド空間を学習空間とし、後に推薦学習部で利用する。

### 2.2 推薦学習部構成

推薦学習部は、学習環境から最適なアイテムをユーザーに推薦する部分である。ユーザーからの推薦依頼がある度に学習、推薦を行う。まず、ユーザーからの推薦依頼を受け、推薦履歴を用いて学習環境上での各エージェントの初期位置を決定するための確率分布を算出する。次に、特徴抽出部で生成された学習環境に、確率分布を元にエージェントを配置する。エージェントは初期位置から意図推定法を用いた強化学習を開始し、一定エ

ピソードが終了するまで学習を行う。ここで、エージェントはユーザーの嗜好情報を利用する。嗜好情報は嗜好学習部で更新される各ゴールに対応した値の集合である。最終エピソードにて到達したゴールに対応するアイテムをそのエージェントが推薦するアイテムとし、ユーザーへの提示を行う。推薦アイテムの決定に嗜好情報を用いるため、意図推定法の任意の関数  $e()$  は式 (1) とする。よって、目標値  $GV$  の更新式は式 (2) となる。

$$e(\vec{GP}_{t,g}) = (1 - \mu)e_1(\vec{GP}_{t,g}) + \mu \cdot e_2(Pre_{t,g}), \quad (1)$$

$$GV_{t,g} = (1 - \eta)GV_{t-1,g} + \eta \cdot ((1 - \mu)e_1(\vec{GP}_{t,g}) + \mu \cdot e_2(Pre_{t,g})), \quad (2)$$

ここで、 $\mu$  は、優先度を用いた評価と嗜好データを用いた評価、それぞれをどの程度重要であるか考えるかの割合である。特に、どちらが重要ということがなければ、通常は 0.5 でよい。 $Pre_{t,g}$  は時刻  $t$  におけるアイテム  $g$  の嗜好データ、 $e_1$  は優先度  $GP$  に対する任意の関数、 $e_2$  は嗜好に関する任意の関数である。ただし、 $0 \leq e_1(\vec{GP}_{t,g}) \leq 1$ 、 $0 \leq e_2(Pre_{t,g}) \leq 1$  である。意図推定法の各エージェント同士の”協調”を決める要素である任意の評価関数  $e()$  を式 (1) とし、優先度に対する任意の関数  $e_1$ 、嗜好に関する任意の関数  $e_2$  を変更することで、様々な推薦問題に対応できると考えられる。

### 2.3 嗜好学習部構成

嗜好学習部は、推薦したアイテムに対するユーザーの評価からユーザーの嗜好を学習する部分である。嗜好学習部では、多層パーセプトロンを用いて各アイテムの嗜好情報を学習、保持する。初期段階で、ユーザーの嗜好情報が無ければ全て初期値であるが、システムが利用されるたびにユーザーの嗜好をオンライン学習する。つまり、多層パーセプトロンの出力が直接ユーザーの嗜好に当たる。

## 3 献立推薦システム

提案システムの有用性を検証するため、実用例の一つとして献立推薦システムを取り上げる。本システムは、ユーザーの嗜好、栄養情報、献立の前後関係を考慮しつつ一定期間の食事の献立を推薦する。栄養情報として、農林水産省が提唱する食事バランスガイドを利用する [11]。食事バランスガイドは、栄養をたんぱく質や炭水化物などの細かい情報ではなく、よりユーザーにとって理解しやすい主食、副菜などの区分で表したものである。献立推薦システムではこの栄養区分を用いて、農林水産省の示す一日に必要なエネルギー量と摂取量を基に、7日間の夕食における栄養の目標値を設定し、その目標値に対して最適な献立の組み合わせを推薦する。

評価関数  $e()$  中の二つの評価関数  $e_1()$ 、 $e_2()$  はそれぞれ式 (3)、式 (4) とする。式 (3) は全エージェントの栄養量の総和を目標値へと最適化し、式 (4) は嗜好データをそのまま使用している。

$$e_1(GP_g) = E_{a \in A_O} [GP_{g,a}] \cdot \exp\left(-\frac{\|\vec{E} - \vec{T}\|}{2\sigma^2}\right) \quad (3)$$

$$\vec{E} = \sum_{a \in A_O} (\vec{E}_a) + \vec{C}_g$$

$$\vec{E}_a = \frac{\sum_{g \in G} (GP_{g,a} \cdot \vec{C}_g)}{\sum_{g \in G} GP_{g,a}}$$

$$e_2(Pre_g) = Pre_g \quad (4)$$

ここで、 $G$  は全ゴールの集合、 $A_O$  は他エージェントの集合、 $a$

は他エージェントの集合のある要素,  $GP_g$  はゴール  $g$  に対する全エージェントの優先度,  $\vec{T}$  は栄養量の目標値ベクトル,  $\vec{E}$  は全エージェントが得る栄養量の期待値ベクトル,  $\sigma$  は分散,  $\vec{E}_a$  はエージェント  $a$  の得る栄養量の期待値ベクトル,  $\vec{C}_g$  はゴール  $g$  の持つ  $N$  次元の特徴ベクトルである。

式 (3) より, 第 1 項は各エージェントが同じゴールへ向かいづらくなるような評価を行う部分, 第 2 項は栄養量を最適化するための評価部である。本システムでは推薦によって同じ献立が選ばれることを許容しているため, 第 1 項により多様な献立を推薦できるよう設計している。各エージェントは全体の栄養量を最適化しながらユーザの嗜好を考慮し, 最適な献立を推薦する。

## 4 評価シミュレーション

### 4.1 問題設定

本シミュレーションでは 7 日間の夕食の献立を推薦する。献立とは複数の料理からなるセットメニューである。また, 本シミュレーションではユーザの評価が得られないため, 全ての献立に対する嗜好情報は等しいものとする。さらに, 推薦履歴も存在しないため, エージェントの初期配置は一様乱数で決定するものとする。

学習環境として用いる二次元グリッド空間は  $20 \times 20$ , エピソード数は 3000, エージェント数 7, ゴール数 40 とし, シミュレーションを行う。意図推定法における他のパラメータは, 著者らの論文と同様の設定を用いる [3]。また, 栄養量を考慮して推薦を行うためには, その目標値が必要である。本シミュレーションで使用する目標値は主食 14.0, 副菜 16.0, 主菜 14.0 とする。この値は, 食事バランスガイドが示す一日に必要な栄養摂取量から朝食, 昼食の栄養摂取量を差引き, 中央値を取ったものである。

また, 推薦に使用する献立データベースは, クックパッド株式会社 が国立情報学研究所と協力して提供しているクックパッドデータセット\*1を元に作成する。献立は複数の料理を含んだ 1 食を表しており, これを推薦候補の献立データとして使用する。

本シミュレーションの評価として, 7 日分の夕食の推薦履歴に対して推薦された献立が栄養摂取量の目標値と近い値となっているかを評価する。

### 4.2 シミュレーション結果

計算機シミュレーションにおいて, 推薦を 100 回行った場合の推薦アイテムの栄養量の平均と標準偏差を表 1 に示す。

表 1 シミュレーション結果

	主食	副菜	主菜
目標値	14.0	16.0	14.0
平均値	14.1	16.8	16.2
標準偏差	0.8	4.2	3.8

表 1 より, 推薦された献立の栄養量が多少の誤差はあるものの目標値と近い値になっていることが確認できる。また, 推薦候補のアイテムが 100 回の推薦中に推薦された回数を図 2 に示す。推薦候補は 40 アイテムであるが, その中に栄養量が完全に等しいアイテムが複数存在するため, それらのアイテムを一つの項目にまとめている。

### 4.3 考察

シミュレーション結果より, 提案システムがユーザへの献立の推薦を行えていることが確認できる。表 1 の目標値と平均値を確認すると, その差が 1~2 程度もしくはそれ以下であるこ

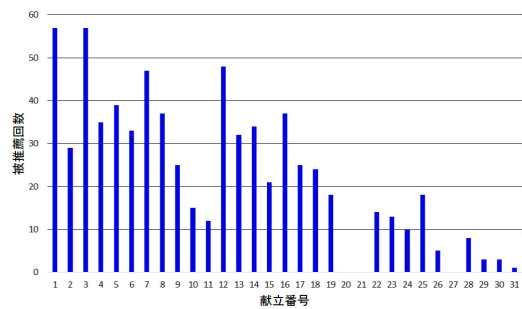


図 2 推薦候補アイテムの被推薦回数

とがわかる。食事バランスガイドでは, 理想的な栄養摂取量は範囲で定められている。それは 7 日間で, 主食 10.5~17.5, 副菜 14.0~17.5, 主菜 10.5~17.5 である。シミュレーション結果の表 1 より, 平均値は理想的な栄養摂取量の範囲に含まれているため, 栄養量の観点から推薦に成功していると考えられる。ただし, 結果にはばらつきがあり, 標準偏差がある程度存在している。

また, 献立データとして使用したデータセットの献立は一般ユーザが作成したものであるため, 栄養量に偏りのあるものが多い。これらは目標値と比較して非常に大きい値であり, シミュレーションの困難さを証明すると共に, 提案システムが機能していることを証明している。

また, 図 2 より, ほとんどの献立が少なくとも一回以上推薦されていることが確認できる。被推薦回数を確認すると, 献立番号の若い献立が推薦されやすくなっていることが見てとれる。この図では副菜の栄養量が少ない献立から昇順に並んでおり, 副菜の栄養量が同じ場合は主菜, 次に主食の順に少ないものから順に並んでいる。そのため, 推薦候補である献立データが全体的に理想的な栄養量よりも多くの栄養量を持つものが数多く存在していることが要因であると考えられる。それらの栄養量の多い献立を選択した場合, 他の献立はどうしても栄養量の少ないものを選ばざるを得ない。そのため, 栄養量の少ない献立番号の若い献立が推薦されやすい傾向があるのだと考えられる。

## 5 まとめと今後の展開

本論文では, 著者らによって提案されている意図推定法を推薦問題へ適用し, 推薦システムを提案することで意図推定法の汎用性を検証した。また, 提案システムを実問題である献立推薦問題へと適用し, 有用性の検証を行った。今後の展開として, 実ユーザの嗜好を学習する部分や推薦履歴から現在の推薦アイテムに影響を与える部分の評価実施が挙げられる。

## 参考文献

- [1] R. S. Sutton and A. G. Barto (三上 貞芳, 皆川 雅章 共訳): 「強化学習」, 森北出版 (2000)
- [2] C. J. C. H. Watkins and P. Dayan: "Q-learning", Machine Learning, Vol.8, No.3-4, pp.279-292 (1992)
- [3] 椿本 樹矢, 小林 邦和: 「意図推定法を用いたマルチエージェント強化学習システムにおける協調行動の獲得」, 電気学会論文誌 C, Vol.134, No.3, pp.117-122(2014)
- [4] 神島 敏弘: 「推薦システムのアルゴリズム (1)」, 人工知能学会誌, Vol.22, No.6, pp.826-837(2007)
- [5] 神島 敏弘: 「推薦システムのアルゴリズム (2)」, 人工知能学会誌, Vol.23, No.1, pp.89-103(2008)
- [6] 神島 敏弘: 「推薦システムのアルゴリズム (3)」, 人工知能学会誌, Vol.23, No.2, pp.248-263(2008)
- [7] 吉井 和佳, 後藤 真孝: 「音楽推薦システム」, 情報処理, Vol.50, No.8, pp.751-755(2009)
- [8] 加島 智子, 折戸 由希子, 山本 久志: 「多期間献立計画問題に対する食育評価モデルの提案と分布推定アルゴリズムによる最適化」, 電子学会論文誌 C, Vol.133, No.8, pp.1576-1585(2013)
- [9] 林 貴宏, 尾内 理紀夫: 「Web 上のレビューを利用した映画推薦システム」, 人工知能学会論文誌, Vol.30, No.1 pp.102-111(2015)
- [10] Teuvo Kohonen: "The Self-Organizing Map", Proceeding of IEEE, Vol.78, No.9, pp.1464-1480 (1990)
- [11] 農林水産省 食事バランスガイド, [http://www.maff.go.jp/j/syokuiku/zissen\\_navi/balance/](http://www.maff.go.jp/j/syokuiku/zissen_navi/balance/), 2016/01/24 アクセス

\*1 情報学研究データリポジトリ クックパッドデータセット:<http://www.nii.ac.jp/dsc/idr/cookpad/cookpad.html>