

議論に基づく推薦システムの構築

～Google 検索エンジンのユーザーインターフェースへの応用～

納田 憲一¹若木 利子²

芝浦工業大学 システム工学部 電子情報システム学科*

1 はじめに

今日、インターネットと WWW の普及により、種々の「Web ページ推薦システム」が開発されてきている。Web ページ推薦システムとは、ユーザーにとって有用な Web ページを選び出し、その情報をユーザーの利用目的にあわせて提供するシステムである。他方、近年、コンピュータネットワーク上のマルチエージェント（ソフトウェア）に人間が行うような議論をシミュレートさせるために、議論の意味論の提案や種々の議論エージェントシステムの開発が進展している。最近、Chesnevar らはこのような議論（argumentation）の技術の応用として、Google 等の検索エンジンの提供する URL の推薦リストに対して、あらかじめ入力された各ユーザーの嗜好データや知識（プリファレンス）を用いて議論の推論を行い、よりユーザーの好みを反映して更新した推薦リストを提示する方法を提案した [1]。しかし彼らは自らが提案する DeLP の言語と議論の意味論を用い、推薦システムの更新も、4 分類（正当化、却下、未定、不明）に基づく。

本研究では、彼らのアプローチに近年、世界的に標準的な意味論として認知された Dung の種々の意味論を適用し、sceptical/credulous な正当化の概念に基づいて 6 段階の分類法を提案する。これにより異なる意味論毎に、先行研究よりユーザの好みをより反映した結果を得ることができた。また、この提案手法に基づいて、議論判定を行いその結果より推薦リストの更新を行うシステムを実装し正当性の検証を行った。

2 準備

2.1 Dung 意味論に基づく議論

Dung が提案した議論の意味論に基づき優先関係の概念を取り入れた議論の意味論を以下に示す。

定義 1 (拡張論理プログラム ELP からなる論証と攻撃関係) [2] ルール集合である拡張論理プログラム (ELP) を P としたとき、 P のルール系列として構成される極小論証の集合を Arg_{SP} 、攻撃関係を $R \subseteq Arg_{SP} \times Arg_{SP}$ として定義する。

定義 2 (ルール間優先関係と論証間優先関係)

P 上の優先関係 (プリファレンス) を決めるものとしてルール間優先関係 $\leq$ を定義し、 $\leq$ が与えられたとき、 Arg_{SP} 上の優先関係として論証間優先関係 $\sqsubseteq$ を文献 [4] に従い定義する。ここで $A_1 \sqsubseteq A_2$ は、「論証 A_1 より論証 A_2 を優先する。」ということを意味する。

定義 3 (議論フレームワーク)

所与の ELP P と P 上のルール間優先関係 $\leq$ より議論フレームワーク AF を、 $AF \stackrel{def}{=} (Arg_{SP}, Def)$ として定義する。Def は「優先関係を考慮した攻撃関係」で、 $Def \stackrel{def}{=} R - \sqsubseteq$ で定義する。

定義 4 (議論の意味論) [3] $S \subseteq Arg_{SP}$ を conflict free な論証集合、単調関数 $F : 2^{Arg_{SP}} \rightarrow 2^{Arg_{SP}}$ を $F(S) = \{A \mid \text{論証 } A \text{ は } S \text{ により defend される}\}$ と定義したとき、議論の各意味論 (semantics) は以下で定義される extension E で与えられる。

- E が complete extension iff $E = F(E)$.
- E が preferred extension (grounded extension) iff E は $\sqsubseteq$ に関する極大 (極小) な complete extension.
- E が stable extension iff preferred extension なる E について、 $Arg_{SP} \setminus E$ の任意の論証を攻撃する論証が E に存在。

2.2 Reinstatement Labellings

Caminada の Reinstatement Labellings の方法を用いて、Dung 意味論が計算できることを以下に示す。[3]

定義 5 (AF-labelling) 抽象議論フレームワーク $AF = (Ar, def)$ の AF-labelling は (全) 関数 $\mathcal{L} : Ar \rightarrow \{in, out, undec\}$ である。これに対して $in(\mathcal{L})$, $out(\mathcal{L})$, $undec(\mathcal{L})$ を次のように定義する。

- $in(\mathcal{L}) \stackrel{def}{=} \{a \in Ar \mid \mathcal{L}(a) = in\}$
- $out(\mathcal{L}) \stackrel{def}{=} \{a \in Ar \mid \mathcal{L}(a) = out\}$
- $undec(\mathcal{L}) \stackrel{def}{=} \{a \in Ar \mid \mathcal{L}(a) = undec\}$

定義 6 (Reinstatement Labellings) AF-labelling: $\mathcal{L}$ が次の条件を満たす時、reinstatement labelling と言う。

- $\forall a \in Ar : (\mathcal{L}(a) = out \equiv \exists b \in Ar : (b \text{ def } a \wedge \mathcal{L}(b) = in))$
 - $\forall a \in Ar : (\mathcal{L}(a) = in \equiv \forall b \in Ar : (b \text{ def } a \supset \mathcal{L}(b) = out))$
- (Ar, def) の任意の reinstatement labelling $\mathcal{L}$ と、各意味論の extension E について、以下の結果が得られる。
- complete extension $E = \{a \in Ar \mid \mathcal{L}(a) = in\}$
 - preferred (grounded) extension $E = \{a \in Ar \mid a \in in(\mathcal{L}) \text{ 但し、} \mathcal{L} \text{ は } in(\mathcal{L}) \text{ が極大 (極小) なラベリング}\}$
 - stable (resp. semi-stable) extension $E = \{a \in Ar \mid a \in in(\mathcal{L}) \text{ 但し、} \mathcal{L} \text{ は } undec(\mathcal{L}) = \phi \text{ (resp. } undec(\mathcal{L}) \text{ が極小)}\}$

定義 7 (Sceptical and Credulous Query-Answering)

- 所与の議論フレームワーク (Ar, def) について、
- 論証 $a \in Ar$ が所与の意味論の下で sceptical (credulous) に正当化される iff 所与の意味論の全ての (ある) $\mathcal{L}$ の集合のラベリングにおいて、 $\mathcal{L}(a) = in$ なる $a \in Ar$ が存在。
 - 論証 $a \in Ar$ が所与の意味論の下で sceptical (credulous) に却下される iff 所与の意味論の全ての (ある) $\mathcal{L}$ の集合のラベリングにおいて、 $\mathcal{L}(a) = out$ なる $a \in Ar$ が存在。
 - それ以外の場合、論証 $a \in Ar$ は所与の意味論の下で sceptical (credulous) に未定。

3 提案手法に基づく推薦システム

3.1 推薦システムの基本構造

提案する推薦システムの構成は先行研究 [1] と同様、以下の 2 つで構成される。(図 1)

1. ユーザーインターフェースを介し検索ワードを受け取り推薦リスト $L = [s_1, s_2, \dots]$ (但し s_i は URL) を出力する Google 等の検索エンジン。
2. 1 から受け取った L を、ユーザーの嗜好情報と L から得られた情報による議論の結果より判定し並び替えた更新推薦リスト L_{Revise} を出力する議論システム。

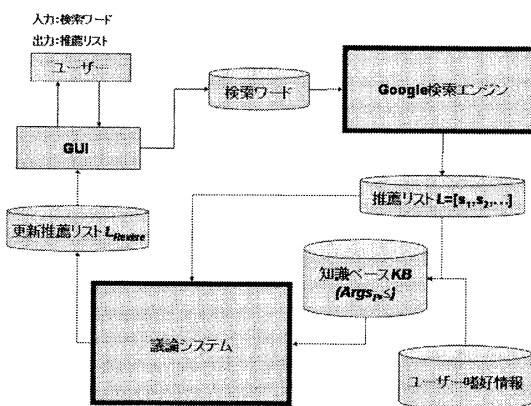


図 1: システム構成図

The recommender system based on argumentation - Application of the user interface of the google searching engine

¹Kenichi Nouda, ²Toshiko Wakaki

*Shibaura Institute of Technology

3.2 提案手法

本研究では、先行研究 [1] で用いられた *sceptical* な議論判定の意味論に対応する *Grounded semantics* だけでなく、*credulous* な議論判定を取り入れた他の *Dung* 意味論を用いて、よりユーザーの嗜好を細かに反映させる。そのため、前述 2 の議論システムを、*Dung* の議論の意味論に基づく本研究で提案する分類方法により実現している。その意味論の計算には文献 [3] の Caminada の Reinstatement Labellings と解集合プログラミングを用いる。

まず、議論の判定結果を各推薦要素の分類に結びつける必要があるため知識ベースの知識を以下のように表現する。

(知識表現)

入力となるユーザーの嗜好情報と推薦リストから得られる事実に情報を知識ベース $KB = (P, \leq)$ として表現する。「推薦リスト内の情報 $s_i \in L$ は適切な情報である」を $rel(s_i)$ で表す (但し、 rel は述語記号)。また、 $s_i \in L$ について $rel(s_i)$ 及び、 $\neg rel(s_i)$ を結論とする論証をそれぞれ Ag, Ng で表す。

さて、*Dung* 意味論に基づき、ユーザーの嗜好に関する *sceptical/credulous* な正当化に基づきより細かな判定結果が得られる。これにより、その判定結果に対応して、推薦の程度の段階を [1] の 4 分類から、以下のように 6 分類により細分化できる。

(推薦クラス)

推薦を 6 段階とし、6 つの推薦クラス $C_1 \sim C_6$ (推薦の強さ $C_1 > C_2 > C_3 > C_4 > C_5 > C_6$) を定義する。各クラスの意味は指定の意味論における *extension* E により以下のように定義する。

- C_1 : $s_i \in C_1$ が全ての E において適切となる。
- C_2 : $s_i \in C_2$ が適切となる E が不適切となる E より多い。
- C_3 : $s_i \in C_3$ が適切か否かが未定である。
- C_4 : $s_i \in C_4$ が適切となる E が不適切となる E より少ない。
- C_5 : $s_i \in C_5$ が全ての E において不適切となる。
- C_6 : $s_i \in C_6$ がユーザーの嗜好に対し無関係である。

[1] における 4 分類は C_1, C_3, C_5, C_6 の 4 つに対応する。

以上のことから、各 $s_i \in L$ を、それぞれ対応する Ag, Ng の所与の意味論における正当化の議論判定より以下のような方法で推薦程度の分類をする。

(議論判定による分類)

対応する Ag, Ng の議論判定 (*sceptical/credulous* に正当化/却下/未定) から以下のように推薦クラスに分類する。(図 2) 但し、 $E \ni Ag$ となる *extension* E の集合を $Ext(Ag)$ として表現し、その数を $|Ext(Ag)|$ とする。これを用い Ag と Ng がともに *credulous* に正当化の場合にエクステンションの数による比較を行う。

1. Ag, Ng 両方存在するとき
2. Ag, Ng どちらか一方のみ存在するとき
3. どちらも存在しないとき

Ag	Ng	sceptical に正当化	sceptical に正当化でない	Ag(Ng) の議論判定	推薦クラス
sceptical に正当化				sceptical に正当化	$C_1 (C_5)$
sceptical に正当化でない				credulous に正当化	$C_2 (C_4)$
				未定	C_3
				credulous に却下	$C_4 (C_2)$
				sceptical に却下	$C_5 (C_1)$
				※括弧内は Ng の場合 の推薦クラスである。	

(注1) エクステンションの数による比較を以下のように行う

エクステンションの大小関係	推薦クラス
$ Ext(Ag) > Ext(Ng) $	C_2
$ Ext(Ag) = Ext(Ng) $	C_3
$ Ext(Ag) < Ext(Ng) $	C_4

図 2: 本研究における分類法

これらのことを踏まえて提案する議論システムの手順は以下のようになる。以下で、論証 A の結論を $conc(A)$ で表す。

(手続き) Arg_Recommender

入力: 知識ベース $(P, \leq)$, 推薦リスト L 出力: 更新推薦リスト L_{Revise}
 Step1: KB より議論フレームワーク $AF = (Argsp, Def)$ を作成。
 Step2: $\forall s_i \in L$ について、 $conc(Ag) = rel(s_i)$ なる $Ag \in Argsp$, $conc(Ng) = \neg rel(s_i)$ なる $Ng \in Argsp$ のそれぞれが *sceptical/credulous* に正当化/却下/未定のいずれに判定されるかを Reinstatement Labellings を用いて計算する。
 Step3: Ag, Ng の判定結果より上記の分類法に基づいて各 s_i を対応する推薦クラスに分類。
 Step4: s_i を推薦の強さの高いクラス順に並び替え L_{Revise} を作成し、これを出力する。

例 (ジャーナリストの記事検索)[1]

ある女性ジャーナリストが鳥インフルエンザに関する記事を探していて、記事の推薦リスト $L = [s_1, s_2, s_3, s_4, s_5, s_6]$ が得られたとする (L の詳細については [1] 参照) ジャーナリストの記事に関する嗜好情報 L から以下のような知識ベース $KB (ELP, P, ルール間優先関係 \leq)$ が得られた。この L の Preferred semantics にお

ける議論による更新を行う。

```

P r1: rel(X) ← author(X, A), trust(A), not ¬rel(X)
/*信頼できる著者 A の書いた記事 X は適切であると考え。*/
r2: ¬ rel(X) ← author(X, A), trust(A),
    outdated(X), not rel(X)
/*古い記事 X は適切でないと考え。*/
r3: trust(A) ← not faked_news(A), not ¬ trust(X)
/*記事の捏造をしていないだろうことから著者 A は信頼できる。*/
r4: ¬ rel(X) ← address(X, Url), biased(Url), not rel(X)
/*この Url の記事 X は意見が偏っているため適切でない。*/
r5: biased(Url) ← thailandian(Url), not ¬ biased(Url)
/*thailand の Url の記事は偏っているといえる。*/
r6: faked_news(chin-yao.lin) ←
    chin-yao.lin は記事の捏造をしている。*/
r7: ¬ faked_news(chin-yao.lin) ← author(X, chin-yao.lin),
    date(X, D), getdate(20090110), (20090110 - D) < 30
/*ごく最近の記事であれば chin-yao.lin は記事の捏造はしていない。*/
...
```

$\leq = \{(r_1, r_2), \dots\}$

$L = \{s_1 : author(s_1, chin-yao.lin),$
 $address(s_1, thai-news.th/...),$
 $date(s_1, 20090103), s_2, \dots\}$

誌面の都合上、記事情報 s_1 についてのみ解説する。KB より議論フレームワーク AF を作成し、以下のような論証が作成できた。

- s_1 の著者は記事の捏造をしていない」という結論を持つ論証
- s_1 の著者は記事の捏造をしている」という結論を持つ論証
- 「信頼できる著者の記事ならば、 s_1 は適切な情報である」という結論を持つ論証 $\rightarrow Ag$
- s_1 は適切な情報でない」という結論を持つ論証 $\rightarrow Ng$

この AF をグラフ表現し得られたラベリング L を以下に示す。(図 3)

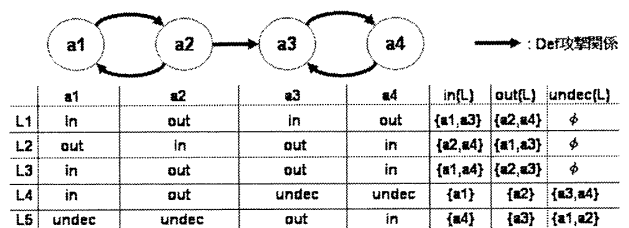


図 3: AF のグラフ表現とラベリング L

得られた $L_1 \sim L_5$ において $in(L)$ が極大となる L_1, L_2, L_3 より以下の Preferred extensions E_1, E_2, E_3 が得られる。

$E_1 = \{a1, a3\}, E_2 = \{a2, a4\}, E_3 = \{a1, a4\}$

この結果より、 $Ag (=a3), Ng (=a4)$ の議論判定はともに *credulous* に正当化となるので、エクステンション数を比較すると、 $|Ext(Ag)|=1, |Ext(Ng)|=2$ で $|Ext(Ag)| < |Ext(Ng)|$ となるので記事情報 s_1 の推薦クラスは C_4 となる。

その他の記事情報についても同様に求め、推薦クラス $C_1 = \{s_3, s_4\}, C_2 = \{s_6\}, C_3 = \{s_5\}, C_4 = \{s_1\}, C_5 = \{s_2\}, C_6 = \phi$ となり、 $L_{revised} = \{s_3, s_4, s_6, s_5, s_1, s_2\}$ が得られる。

また、grounded semantics の場合の結果は $C_1 = \{s_3, s_4\}, C_2 = \phi, C_3 = \{s_1, s_5, s_6\}, C_4 = \phi, C_5 = \{s_2\}, C_6 = \phi$ によって $L_{revised} = \{s_3, s_4, s_1, s_5, s_6, s_2\}$ となり先行研究 [1] と一致する結果が得られた。

3.3 実装・検証・評価

提案手法によるシステムを JAVA 言語によって実装した。本システムは、知識ベース KB をテキストファイルで記述し、入力として意味論 (complete, preferred 等) を指定することで実行する。出力として更新された推薦リスト $L_{revised}$ を画面表示しテキストファイルとして出力している。種々の例題について検証を行い、手計算と同様の結果を得られた。Grounded semantics の場合先行研究と一致する結果を得られたことで提案手法の正当性を検証できた。

4 おわりに

本システムは汎用性があり、検索エンジンの提供する URL 推薦リストの更新のみならず他のアプリケーションへの応用と評価が今後の課題である。

参考文献

- [1] C. Chesnevar, A. Maguitman, G. Simari: Argument-Based Critics and Recommenders: A Qualitative Perspective on User Support Systems: Data & Knowledge Engineering 59, page 293-319, 2006
- [2] R. Schweinmeier, M. Schroeder: A Parameterised Hierarchy of Argumentation Semantics for Extended Logic Programming and its Application to the Well-founded Semantics: Theory and Practice of Logic Programming, pages 207-242, 2005
- [3] T. Wakaki and N. Nitta, Computing Argumentation Semantics in Answer Set Programming, Proc. 2nd International Workshop on Juris-Informatics (JURISIN2008), pp. 32-41, 2008.
- [4] 関口 知之, 若木 利子: 動的ブリアレンスを扱う議論エージェントシステムの構築: 電子情報通信学会研究報告 (信学技報), AI2008-3, pages 15 - 20, 2008