

ビヘイビアベースの情報組織化に関する研究

下田 修[†] 嘉数 侑昇[‡]

北海道大学大学院情報科学研究科[†] 北海道情報大学[‡]

1 はじめに

今日のインターネットの普及により、人々は膨大な量の情報に容易にアクセスすることができるようになった。また、Wiki やブログ等の普及とともに、日々感じた雑多なことを日記として綴ったり、自分が得た断片的な知識をメモとしてネットワーク上に公開したりする人が増えてきている。更に最近では、これらのツールを企業活動に取り入れ、情報共有や Know-who のための知識基盤として用いる例も出てきている。

一方このような断片的な情報は、公開されているものの一義的には本人のためのものであったり、書籍のような一貫した文脈を持たず、それを読み取るユーザによって初めて文脈的意味や価値が与えられることも多い。

そこで、利用者視点からの情報の活用技術として、キーワードによる単純な文書検索だけでなく、情報の分類・抽出・組織化などの技術が求められており、様々な研究がなされている[1]。

このうち、情報組織化の分野では、情報の関連性を扱う技術を応用し、知識共有や発想支援等の研究が行われている(例えば[2],[3])。これら既存の多くの研究では、人間の発言からキーワードを抽出してキーワードマップとして可視化したり、連想辞書を用いて関係するキーワードを表示することで人間に対して思考の深化や気づきを促すものが多く、これらは、新しい発想を得ようとするブレインストーミングなどにおいて、議論の幅を広げることができ有用であるとされる。

しかしながら、誰も話題にしなかった内容に対しては基となるキーワードが与えられないため、気づきの促し方は「キーワードがない(すなわち、まだ議論されていない)空白部分がある」という受動的な促し方にとどまってしまう。しかも、その空間にどの程度の「隙間」があるのかもわからない。

人間が集団で議論すると新しい考えを得やすい理由の1つは、各個人が異なった背景知識や物の見方を持っており、与えられたテーマに対して様

々な意見を出しあうことで各自が自分と異なる視点からの意見を聞けるからであると考えられる。従って、「関連知識」だけでなく、「異なる視点からの知識」を提示することでより多様な情報の組織化を促すことが出来ると言える。

以上の議論に基づき、与えられた話題に対してより多くの視点からの情報を出すようなエージェントのビヘイビアに注目することで多様な情報組織化をもたらし、より広範に話題空間をカバーするように情報を提示可能なエージェントモデルに関する研究を行なう。

2 発散的思考支援モデル

以下では、本モデルの主要な要素である言語とエージェントのモデル設計について述べる。

2.1 言語設計

本システムでは、情報の提示・組織化等は会話をベースとして行う。ここでは、会話に用いる言語は以下のように定式化される。

L : Language

$$w = \{\alpha, \beta, \gamma, \dots\} \in L, w: \text{alphabet of } L \dots (1)$$

$$w \supseteq (w_i, w_j, \dots, w_k) \dots (2)$$

$$w_i R w_j = r \text{ where } 0 \leq r < 1, i \neq j \dots (3)$$

, R : Relationship between w_i and w_j

ここで、(2)式は、ある単語 w が他の単語集合のラベルとして機能することを意味する。また、式からわかるように、この集合関係は入れ子にすることができる。これを連想辞書として用いれば、周辺知識を提示することも可能となる。

次に、(3)式は、ある単語が他の単語との関連度を持つことを意味する。具体的には、正規化した共起頻度を値などとして定義される。

$$s = \bigcup_i s_i \quad : \text{sentence} \quad \dots (4)$$

$$s_i = \bigcup_j w_j \quad : \text{element of sentence} \quad \dots (5)$$

s は1つの発言(文)を、 s_i は文素を表す。エージェントの1回の発話は1つの文からなる。

本モデルでは、(3)式で定義される R を用いて、各単語の内部表現に、単語自体を他の単語との関連度を用いて定義するベクトル空間モデルを採用した。

A Study On Information Organization Based On Agent Behavior

[†]Osamu Shimoda, Graduate School of Information Science and Technology, Hokkaido University

[‡]Yukinori Kakazu, Hokkaido Information University

2.2 エージェント設計

以下に、エージェントのビヘイビアのモデルを示す。

$$\text{Behaves } \mathbf{b} = \sum \mathbf{w}_{proj} \rightarrow c \quad \dots(6)$$

$$\text{Subject to } \mathbf{w}_{proj} = \left(b_{w_1 \in T}, b_{w_2 \in T}, \dots, b_{w_k \in T} \right) \mathbf{w} \quad \dots(7)$$

$$\text{where } \mathbf{b}_w = \overrightarrow{gw} = \overrightarrow{ow} - \overrightarrow{og} \quad \dots(8)$$

$$\text{and } \mathbf{g} = \frac{1}{|T \cap L_a|} \sum_{w_i \in T \cap L_a} \mathbf{w}_i \quad \dots(9)$$

ここで、 T はディスカッションのテーマとして与えられた単語の集合を、 L_a はあるエージェントが知っている単語の集合を表す。

(6)式において、ベクトル \mathbf{b} は会話の「バランスベクトル」と呼び、与えられたテーマに対してバランス良く様々な観点から会話が行われているかどうかを表すものである。従って、(6)式において $c = 0$ とすれば、このベクトルのノルムが 0 に近づいているほど偏り無く会話が行われていることを意味する。

機械エージェントは、誰かが発言するたびに内部のバランスベクトルを更新し、傾いたバランスを元に戻すようなキーワードを発言して、会話のバランスを戻そうとする。これにより、テーマに対して触れられていない部分を能動的に提示し、網羅的に議論を進めることができると考えられる。

このとき、バランスベクトルを単純に発言内の単語ベクトルの和としてしまうと、次元ののろいのためノルムを小さくできる適当な単語ベクトルを見つけられないおそれがある。そこで、各単語ベクトルを、テーマとして与えられた単語集合の重心 ((9)式で表される) を原点とした $|T|$ 次元のテーマ空間に射影し、(7)式に従って射影されたベクトルの和をバランスベクトルとした。こうすることで、テーマとして与えられた単語のうちどれに対してより多く発言されているかを知ることができる。

3 実験

提案モデルに基づき、話題空間に対して偏り無く情報を提示できるかどうかを確かめるために、実験を行った。

実験では、エージェントの 1 回の発言は 1 単語で構成し、単語は 0~199 の番号とした。また、全てのエージェントで同じ単語集合を共有することとした。なお、単語ベクトルの要素 (すなわち関連度) は乱数で与えた。実験パラメータは、テーマ単語番号: 23, 4, 10, 単語数: 200, 最初の発言単語: 40, エージェント数: 3 とした。

また、同じパラメータで、知識が分散している場合の実験も行った。

以下に、実験結果を示す。なお、値は全てエージェント 1 のものである。

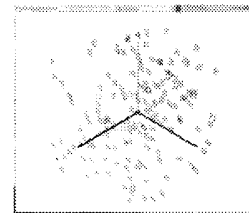
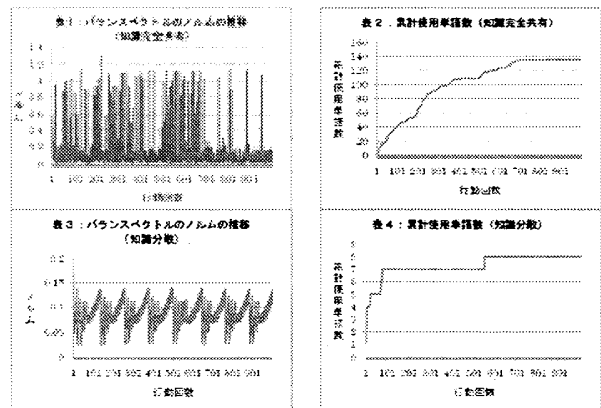


図 1 : テーマ空間に射影した提示単語

4 考察

まず、知識を完全に共有した場合について見てみると、表 1 ではノルムが 1.0 程度で抑えられている。図 1 と合わせて考えて、話題が特定の方向に偏っておらず、様々な情報が提示されていることが確認できた。また、表 2 から分かるように、提示した単語も様々なものが選ばれていた。

一方、知識を分散させた場合では、表 4 を見るとわかるように使用する単語が増えていない。また、表 3 のバランスベクトルのノルムの推移と合わせて考えて、他エージェントの発言に対して適切な単語が見つからず周期的に同じ単語を繰り返し提示していることがわかった。

5 今後の課題

未知の単語を受信した場合に、発言者にその単語の情報を問い合わせる単語を学習する機構を備えること、また、単語の関連度に偏りがある場合について検討することが今後の課題である。

6 参考文献

- [1] 武田英明「ネットワークを利用した知的情報統合」人工知能学会誌, Vol.11, No.5, pp.680-688, 1996
- [2] 平田高志, 村上晴美, 西田豊明「連想表現と分身エージェントを用いたコミュニティにおける知識共有支援」人工知能学会誌, Vol.16, No.2, pp.225-233, 2001
- [3] 西本一志, 角康之, 門林理恵子, 間瀬健二, 中津良平「マルチエージェントによるグループ思考支援」電子情報通信学会論文誌, Vol.J81-D-I, No.5, pp.478-487, 1998