

Talking-永田町 - 政治家エージェントの疑似会話システム -

村上 晴美[†] 平田 高志[‡]

[†] 大阪市立大学学術情報総合センター [‡] 防衛庁陸上自衛隊

1. はじめに

我々はこれまでに、コミュニティにおけるメンバーの相互理解や知識の共有を支援するシステム CoMeMo-Community[1, 2] を研究開発してきた。このシステムではコミュニティのメンバーが連想表現と呼ぶ知識表現を用いて知識を記述し、分身エージェントと呼ぶメンバーの身代りとなるエージェントに与えておく。分身エージェントは仮想会話空間を監視し、自分の連想表現に含まれる単語が選択されると、単語の周囲に近づいてくる。ユーザが分身エージェントを指名すると、分身エージェントの持つ連想表現が展開、表示され、音声でよみあげられる。展開された連想表現の中からユーザが単語を選択すると、再びその単語を持つ分身エージェントが近づいてくる。このような一連の動きを疑似会話とみなしている。疑似会話を通して、ユーザは自分の知識を公開し、他人の知識を理解し、他人の知識に新しい知識をつけ加えて公開する。CoMeMo-Community では、実際に人間が行うような連想的な会話で連想表現と分身エージェントという概念を用いて可能かどうか実験を行い、一定の成果を得た。その理由として、コミュニティではある程度背景知識を共有しているため他人が作成した連想表現を理解できること、他人の連想表現を見ながら自分の連想表現を追加できること、などが考えられた。

本研究では、新聞記事中の人間の発言テキストからエージェントの知識を構築して疑似会話が可能かどうかを検討する。本研究では、(1)「新聞記事中の人間の発言テキストからエージェントの連想表現を作成することにより疑似会話が可能である」、(2)「コンテンツとして政治の話題が有効である」、(3)「新聞記事から作成した疑似会話システムが話題の理解に役立つ」という仮説をたて、CoMeMo-Community をベースとする政治家エージェントの疑似会話システム Talking-永田町を試作した。

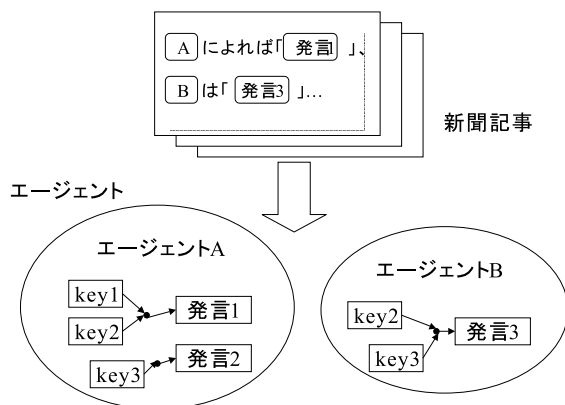
本稿では2. で試作したシステムの概要を述べ、3. で2000年衆議院総選挙を題材として行った実験について述べる。

2. Talking-永田町

Talking-永田町は、政治家エージェントの疑似会話システムである。Talking-永田町は、(1) エージェント知識の構築と、(2) エージェントの疑似会話の2段階から構成される。

図1にシステムの概要を示す。(1) エージェント知識の構築では、新聞記事中の政治家の発言(「」で囲まれた文字列)を手作業で抽出し、政治家毎に、発言に含まれる主要なキーワード群と発言を結びつける連想表現を作成する。(2) エージェントの疑似会話では、構築された政治家エージェントの知識を用いて、入力されたキーワード(key1)をきっかけにエージェントAが発言し、その中に含まれるキーワード(key2)をきっかけにエージェントBが発言し、さらにその中に含まれるキーワード(key3)をきっかけに再びエージェントAが発言する、というように疑似会話を行う。

(1) エージェント知識の構築



(2) エージェントの疑似会話

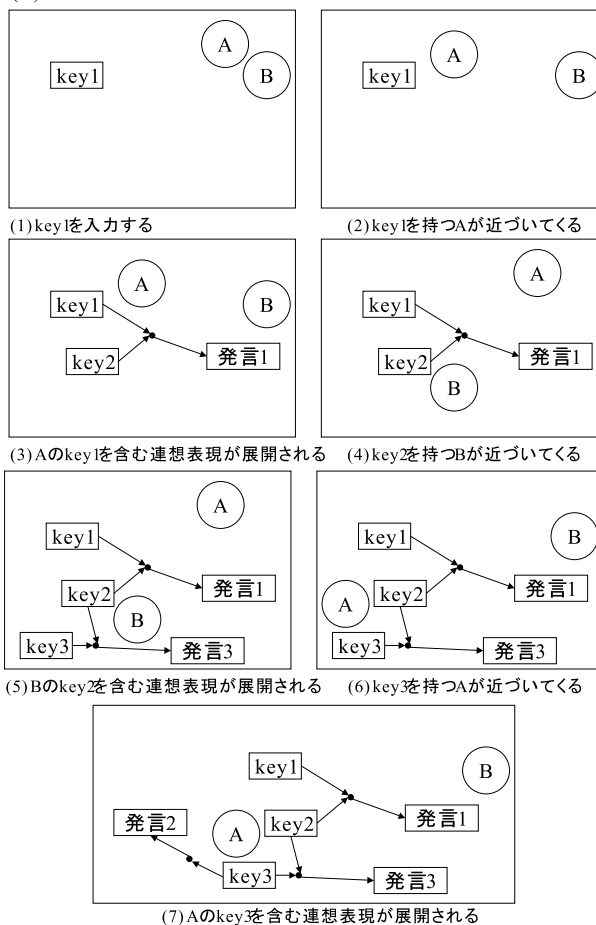
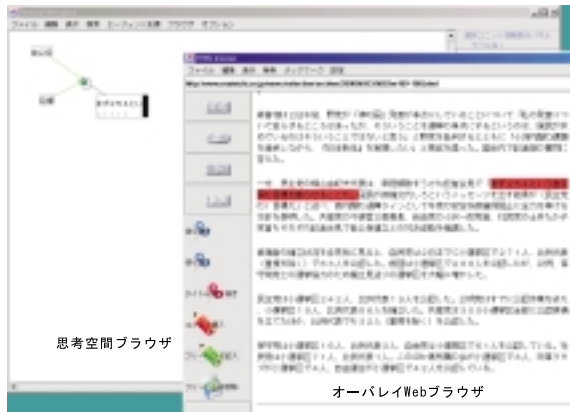


図1: Talking-永田町の概要



(a)Memory-Organizer での発言の切抜き



(b) 作成された政治家エージェントの知識
図 2: エージェント知識の構築

2.1 エージェント知識の構築

エージェント知識の構築には、個人の外化記憶構築システム Memory-Organizer[3, 4] のオーバーレイ Web ブラウザと思考空間ブラウザを利用し、以下の手順で行う。

1. オーバレイ Web ブラウザに表示された新聞記事上で、政治家の発言をマウスでなぞると、なぞられた箇所からキーワードが切り出され、発言と関連づけられて連想表現として思考空間ブラウザ上に表示される。
2. 思考空間ブラウザ上でキーワードの追加、編集、削除を行う。
3. 思考空間ブラウザ上で連想表現をエージェント知識として保存する。

キーワードの切り出しアルゴリズムは、(1) ひらがなと特殊文字以外の文字が 2 つ以上続いた文字列を抽出したあと、(2) 不要語処理を行うというシンプルなアルゴリズムを用いている。詳細は [2, 3] 参照。

図 2 にエージェント知識構築の画面例を示す。たとえば、オーバーレイ Web ブラウザに表示された記事中の鳩山氏の発言「まず 254 という自公保の目標を割らせることだ。」という部分をマウスでなぞると、思考空間ブラウザに「自公保」「目標」というキーワードと記事全体へのリンクが連想表現として作成される(図 2(a))。このようにして作成された連想表現群を政治家エージェントの知識として Talking-永田町に取り込む(図 2(b))。

2.2 エージェントの疑似会話

エージェントの疑似会話は、(1) キーワードの選択と(2) エージェントの指名(選択実行)の 2 段階から成り立っており、これらをユーザが行う手動モード、システムが行う自動モードの 2 種類のモードがある。

手動モード 手動モードでは、以下のようにキーワードの選択とエージェントの指名ができる。

- ユーザがキーワードを選択すると、選択されたキーワードを持つエージェントが近づいてくる。
- ユーザがエージェントを指名すると、発言に含まれるキーワードを含む連想表現が表示され、エージェントが発言する。

自動モード 自動モードでは、システムがどのようなキーワードを選択し、エージェントを指名するかや、どのようなタイミングで切り替えるかなどについて多様な考え方が可能である。

ここでは、「政治家が人の発言の最後の部分に反応して、自分の言いたいことを主張する」状況をシミュレーションする。ある話題(キーワード)をもとに一人の人間が発言し、その発言の中で展開された別の話題(キーワード)をもとに、人間が発言するアルゴリズムの一例を以下に示す。

```

begin
  ユーザが「現在のキーワード」を入力または選択する;
  repeat
    if 「現在のキーワード」がまだ発言されていない連想表現の中に見つかる
    then
      begin
        「現在のキーワード」を持つエージェントが近づいてくる;
        システムがエージェントと連想表現をランダムに選択する;
        選択された連想表現が表示される;
        エージェントが発言する;
        最後に表示されたキーワードが「現在のキーワード」になる;
      end
    else 直前に表示されたキーワードが「現在のキーワード」になる;
  until すべてのキーワードが発言される
end.

```

3. 実験

2000 年衆議院総選挙(2000/6/25)を題材として予備的な実験を行った。

3.1 実験 1

目的 実際の新聞記事からどの程度、政治家のエージェント知識を構築して疑似会話が可能か検討する。

方法 以下の 2 つの情報源を対象としてエージェント知識の構築を行う。

- (a) asahi.com の選挙特集記事¹(以下「asahi.com 選挙特集」)

¹ <http://www.asahi.com/>

エージェント	asahi.com 選挙特集		毎日 6 月	
	発言数	キーワード数	発言数	キーワード数
扇エージェント	11	18	3	7
小沢エージェント	10	22	6	12
土井エージェント	10	15	1	1
鳩山エージェント	10	19	12	20
不破エージェント	10	29	1	2
森エージェント	4	5	21	26
合計	55	108	44	68

(注) キーワード数はエージェント毎に重複を除く

(b) 毎日インタラクティブのニュースセレクションの 6 月 1 日から 25 日までの記事² (以下「毎日 6 月」)

エージェント化する政治家として日本の主要な政党の党首などの扇氏、小沢氏、土井氏、鳩山氏、不破氏、森氏の 6 名を選択した。

手続きは、(1) 上記 6 人の発言を含む衆議院選挙に関連する記事を主観的に選択し、(2) 2. で述べた手法を用いてエージェント知識を構築する。

被験者は、asahi.com 選挙特集は第二著者(当時 30 歳男性)が担当し、毎日インタラクティブ 6 月は大学 3 年生の 19 歳女性が担当した。

結果 表 1 に作成されたエージェント知識の概要を示す。

asahi.com 選挙特集 (SD=2.6) が毎日 6 月 (SD=7.9) よりも発言数の偏りが少ない。通常の記事である毎日 6 月から作成した方は、当時首相の森氏と野党第一党党首の鳩山氏の発言数が他と比べて非常に多いことがわかる。

以下に前者を対象として「景気」というキーワードをきっかけとした手動モードの疑似会話の例を示す。図 3 に画面例を示す。

- 扇エージェント、小沢エージェント、土井エージェント、鳩山エージェント、不破エージェント、森エージェントを呼び出す。
- 「景気」を入力する。
 - 扇エージェント、小沢エージェント、鳩山エージェント、不破エージェント、森エージェントが近づいてくる。
 - 扇エージェントを指名する。
 - 扇エージェント発言「景気はせっかく上向いてきたが、まだ本物の回復にはなっていない。」
 - 不破エージェントを指名する。
 - 「政府」「失業者」「国民」を含む連想表現が表示される。
 - 不破エージェント発言「政府は景気が回復したと大騒ぎしているが、失業者や倒産の状況を見ると、国民には実感が無い。」
 - 鳩山エージェントを指名する。
 - 「政権」を含む連想表現が表示される。
 - 鳩山エージェント発言「財政をよくすれば、景気もよくなるのに、今の政権はなぜそれができないのか。」
 - 小沢エージェントを指名する。
 - 「自民党」「幹事長」「補正予算」「公共事業」を含む連想表現が表示される。
 - 小沢エージェント発言「自民党の幹事長が秋に補正予算を組むというが、従来型の公共事業の追加であり、景気はよくなる。」
- 「公共事業」を選択する。
 - 小沢エージェント、鳩山エージェント、不破エージェントが近づいてくる。

(b) 小沢エージェントを指名する。

- 「金利」を含む連想表現が表示される。
- 小沢エージェント発言「単に金利をいじったり、公共事業にカネを使うだけではよくならない。」

(c) 不破エージェントを指名する。

- 「予算」「ゼネコン」「福祉」「教育」を含む連想表現が表示される。
- 不破エージェント発言「予算をゼネコン中心の公共事業から福祉、教育に切り替えるべきだ。」

自動モードで疑似会話を試行したところ、両者ともにある程度発言が継続して行われた。たとえば、「景気」というキーワードから疑似会話を試行したところ、asahi.com 選挙特集では 22 の発言(全体の 40%)が、毎日 6 月では 30 の発言(全体の 68%)が継続して行われた。

今後は、どのようなテーマに対してどのような新聞記事を対象とするのがよいのか継続して検討する必要がある。また、疑似会話の有効性についてより定性的、定量的な評価実験が必要である。

3.2 実験 2

目的 エージェント知識の構築手法を用いて、政治家の発言の特徴が把握できるか検討する。

方法 エージェント知識に含まれるキーワードを使って、数量化 III 類による 2 次元表示を行う。

以下の 2 種類のキーワード群を対象として、エージェント毎に 1 回以上出現したキーワードを 1、その他を 0 として算出する。

- 表 1 に示す合計キーワード 108 個の中から重複を除く 93 個のキーワード
- 上記 93 個のキーワードの中から、2 つ以上のエージェント知識において出現した 16 キーワード：景気、国民、政治(以上 4 エージェントにおいて出現、以下同様)、公共事業、政権(以上 3 エージェント)、公明、自民、自民党、政党、選挙、増税、不況、平和、保守党、野党、連立政権(以上 2 エージェント)

結果 (a)、(b) ともに、本研究で提案したエージェント知識の構築手法を用いて、数量化 III 類の第 1 軸と第 2 軸を用いたラベル付散布図の出力が可能であった。両者ともに見た目の特徴には大差がなかった。ここでは、(b) を対象として数量化 III 類を用いて、第 1 軸を X 軸、第 2 軸を Y 軸として出力したラベル付散布図を示す(図 4)。

以下に、エージェント毎に 2 回以上出現したキーワードを示す。

- 扇エージェント：野党 (2 回)
- 小沢エージェント：公共事業、国民 (2 回)
- 土井エージェント：沖縄、憲法 (3 回)
- 鳩山エージェント：政治 (4 回)、国民、政権 (3 回)、自民党、勇氣 (2 回)
- 不破エージェント：共産党、国民、政治 (2 回)
- 森エージェント：公明党、日本経済、野党 (2 回)

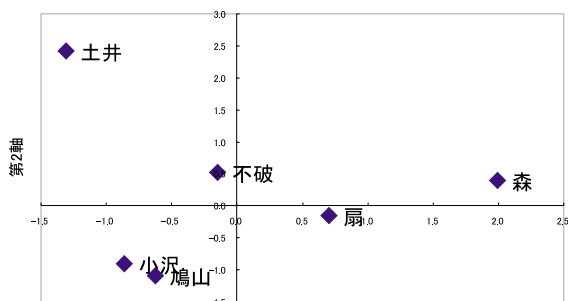
図 4(a) からは、各エージェントが、森エージェント、土井エージェント、小沢エージェントと鳩山エージェントの 3 つのグループに分かれること、扇エージェントと不破エージェントはほぼ中間に位置することがわかる。

ここで扱うデータは発言に出現したキーワードのみであり、そのキーワードに賛成か反対かまでは扱っていない。

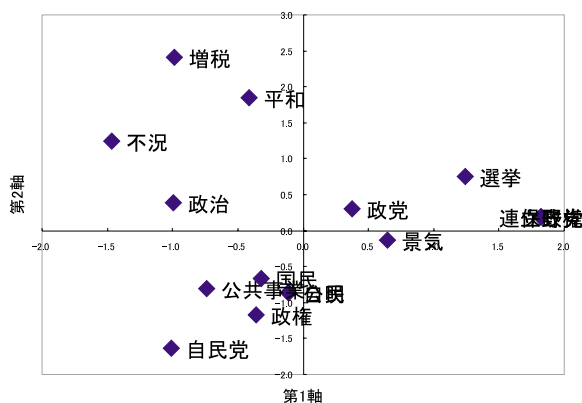
² <http://www.mainichi.co.jp/>



図 3: 例題: 2000 年衆議院総選挙



(a) エージェント



(b) キーワード

図 4: 数量化 III 類によるラベル付散布図

ないため、小沢エージェントに「公共事業」、鳩山エージェントに「自民党」、森エージェントに「野党」が含まれるように、批判する立場のキーワードが頻繁に出現する。土井エージェントに最も多いキーワードは「沖縄」「憲法」であるが、これらは他では全く出現せず、土井エージェントは他エージェントと比べて際だって異なることがわかった。

軸の解釈は困難であるが、本研究で提案した簡単な方法でもある程度政治家エージェントの発言の特徴が把握できることを確認した。もちろんこれらは政治家自身の発言の特徴だけではなく、新聞メディアの傾向を強く反映しているだろう。

4. おわりに

政治家エージェントの疑似会話システム Talking-永田町を試作した。新聞記事中の政治家の発言テキストから手作業で連想表現を作成してエージェントに与える場合に、疑似会話が成立する可能性と、話題の理解に役立つ可能性を示した。

参考文献

- [1] 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の知識共有支援, 人工知能学会論文誌, Vol.16, No.2, pp.225-233, 2001.
- [2] 平田 高志: 外化記憶の構築と共有の支援に関する研究, 奈良先端科学技術大学院大学情報科学研究科博士論文, 2001.
- [3] 村上 晴美, 平田 高志: WWW からの情報の獲得と整理 - 思考・興味空間ブラウザー-, 情処研報, Vol.2001, No.20, pp.167-174, 2001.
- [4] 村上 晴美, 平田 高志: Memory-Organizer: 個人の外化記憶構築システム, 2001 年度人工知能学会全国大会 (第 15 回) 論文集, 3F1-03, 2001.