

ロボット会話への利用を想定した要望に適する場所の提案手法

芋野 美紗子 吉村 枝里子 土屋 誠司 渡部 広一

同志社大学 理工学部, 大学院 工学研究科

{mimono, eyoshimura}@indy.doshisha.ac.jp

{stsuchiy, hwatabe}@mail.doshisha.ac.jp

1 はじめに

ロボットが人間にとって違和感の無い存在となるためには、単に一定の機能を繰り返すような存在ではなく、円滑なコミュニケーションの元で行動を行う必要がある。人間同士のコミュニケーションにおいて最も重要となるのは自然言語による会話であり、ロボットが人間と円滑なコミュニケーションを行う上で、人間らしい会話能力は必須といえる。しかし人間がロボットに対して発話を行い、それに対する決まった返答を行うだけでは人間らしい会話は成立しない。会話中の発話から話題の掘り下げや新たな話題の提供、何かしらの提案を行うなどといった「自発性を感じる返答」を行うことで人間らしさが生まれると考える。

そこで本稿では人間とロボットの円滑なコミュニケーションを担う技術の一端として、なにかしらの要望を表す文に適合する場所語を提案する手法について示す。提案手法により、例えば「温泉に入りたい」といった人間の発話から「由布院はどうですか」といった要望に適する場所を提案する応答を生成したり、「アトラクションに乗りたい」のような発話から「東京ディズニーランドは楽しいですよ!」といった雑談的な応答の生成を行うことができると考える。

2 概要

本手法は、入力された要望を表す文（要望文）に適する場所語を Web 上から獲得してユーザに提案する。図 1 に提案手法の概要を示す。

入力として扱う要望文は、例えば「羊と遊びたい」や「アトラクションに乗りたい」のような希望の助動詞「たい」を含む1文とする。要望文から「羊」や「アトラクション」のような要望の対象となる名詞（要望名詞）および「遊ぶ」「乗る」のような要望の動作を表す動詞（要望動詞）を抽出し、これらの語を用いて

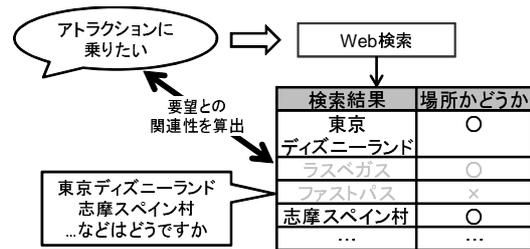


図 1: 提案手法の概要

Web 検索を行う。検索結果から得られる語句について場所語であるかの判断を行い、これらの獲得された語と要望文との関連性を定量的に算出することで要望に適合した場所の提案を行う。

3 語概念連想システム

語概念連想システムとは、人間のように柔軟な語句の意味理解を行うための機構である。語の意味を定義した「概念ベース」および関連性の定量化を行う「関連度計算方式」により語概念連想システムは構築される。本稿ではこの語概念連想システムの機構を用い、「場所語候補が確かに場所を表す語であるかの判断」および「要望文と場所語との関連性の定量的な判断」を行う。

3.1 概念ベース

概念ベース [1] は複数の電子化国語辞書などの見出し語を概念として定義し、人間が持つ概念への常識的な知識をモデル化した知識ベースである。ある概念の意味定義は、属性と呼ばれる他の概念群と属性それぞれの重要さを表す重みによってなされる。概念ベースの具体例を表 1 に示す。

表 1: 概念ベースの具体例

概念	属性
蝸牛	(腹足類,1.62)(螺旋,0.71)(梅雨,0.35)...
梅雨	(紫陽花,1.14)(季語,0.49)(蝸牛,0.48)...

概念ベースにおいて属性に現れる語句は、全て概念として定義されている。そのため、概念「蝸牛」の意味定義を行う属性「梅雨」も概念ベースにおいて意味定義がなされている。このような意味定義の連鎖的な繋がりにより、より人間らしい語句の意味定義が可能となる。

3.2 関連度計算方式

関連度計算方式 [2] は概念と概念の関連性を関連度とよばれる数値で定量的に表現する手法であり、その有効性が示されている。関連度は 0.0 から 1.0 の値を取り、概念間の関連が強いほど大きな値を示す。例えば概念「自動車」と概念「車」の関連度は 0.912 と非常に大きな値を示す。一方、概念「自動車」と概念「学校」の関連度は 0.012 となり、関連が薄い概念同士では小さな値となる。関連度は概念同士の属性の対応により算出される。互いが持つ属性の内、最も意味が近いもの同士の組を作った上でそれぞれの重みを用いて関連度を算出する。

4 提案手法の流れ

図 2 に提案手法の流れを示す。

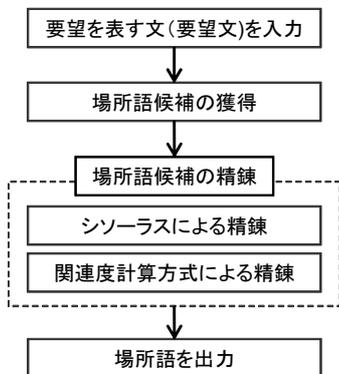


図 2: 提案手法の流れ

まず入力された要望文中の要望名詞および要望動詞

を用いて Web 検索を行い、場所語候補となる語句の獲得を行う。次に、得られた場所語候補についてシソーラス [3] による精練を行う。これは場所語候補として得られた語句のうち、確かに場所を表す語のみを選択するための処理である。次に関連度計算方式による精練によって要望文に適合する場所語の選別を行う。この際、Web を用いた語句の概念化手法 [4] を利用して要望の概念化を行う。これを要望概念とし、各場所語との関連度を算出した上で最終的に関連度が高い場所語を提案として出力する。

5 場所語候補の獲得

場所語候補の獲得では Web 検索により提案する場所語となりえる語句の獲得を行う。Web 検索の入力として、要望名詞・要望動詞・「旅行」・「場所」の&検索を用いる。例えば、「アトラクションに乗りたい」のであれば、「アトラクション&乗る&旅行&場所」が検索の入力となる。この検索結果から得られる Web ページ 100 件に対して形態素解析を行い、tf・idf を用いて重み付けを行う。そして重み上位 30 語を場所語候補として取得する。この時、「数字を含む語句」「助詞を含む語句」「動詞を含む語句」「記号を含む語句」については提案する場所語の候補として適さないと判断し、削除する。図 3 に場所語候補の獲得例を示す。

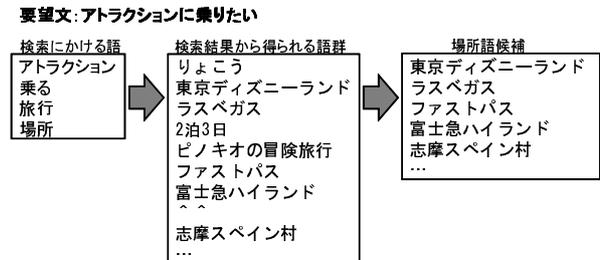


図 3: 場所語候補の獲得例

得られる語群のうち、たとえば「2泊3日」のような語句は数字が含まれているため削除される。

6 場所語候補の精練

獲得した場所語候補について、シソーラスおよび関連度計算方式による精練を行う。場所語候補の語句のうち、確かに場所を表す語句のみを残し、さらに要望文との関連度を算出することで提案に適した語句を抽出する。

6.1 シソーラスによる精練

シソーラスによる精練では、まず得られた場所語候補がシソーラス上に存在しているかを調査する。存在している場合には親ノードを探索し、ノード「地名」が存在する場合には場所を表す語と判断する。シソーラスに場所語候補が存在しない場合には、シソーラスマッピング [5] により場所語候補が所属すべきノード（所属ノード）の候補を取得する。

シソーラスマッピングとは、関連度および Web 検索を用いてシソーラス上に定義されていない語句の所属すべきノードを判別する手法である。Web 検索においては各々のノードに対してあらかじめ割り振った共起しやすい動詞（ノード動詞）を利用する。図 4 にシソーラスの一部を示す。

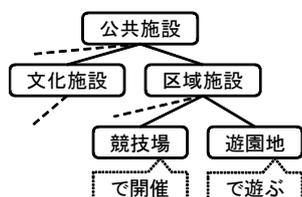


図 4: シソーラスとノード動詞

まずノードと場所語候補の関連度を算出する。さらにそのノードに割り振られているノード動詞と場所語候補を連結した状態の語句での Web 検索 HIT 件数、場所語候補とノードとの AND 検索 HIT 件数を調査する。これらの値を掛け合わせたものを、そのノードに場所語候補が所属する確からしさとし、値の大きい上位 6 件のノード名を所属ノード候補として取得する。

所属ノード候補の上位 6 件のうち、4 件以上が場所語と判断された場合には、場所を表す語として候補に残す。表 2 に場所語「東京ディズニーランド」のシソーラスマッピングにより得られる所属ノード候補および場所を表す語かの判断例を示す。

表 2: シソーラスマッピングによる場所の判断

所属ノード候補	場所かどうか
遊園地	○
宿泊施設	○
動物園・植物園	○
雑誌	×
道路	○
券	×

この例では取得された所属ノード候補 6 件のうち 4 件から場所語と判断されたため、「東京ディズニーランド」は場所を表す語という結果を得る。

6.2 関連度計算方式による精練

関連度計算方式による精練では、Web を用いた語句の概念化手法を利用して要望文を概念化する。これは要望名詞および要望動詞の & 検索により得られた Web ページ群から属性を抽出することで実現している。図 5 に要望概念の属性例を示す。

概念	属性
アトラクション & 乗る	アトラクション, バス, 開園, 絶叫, パーク, 園内, リゾート, パレード, 混雑, 楽しい, コースター, パスポート, 身長, タワー, ジェット, チケット, 恐怖, キャスト, 入園, ゴンドラ, 乗り物, ...

図 5: 要望概念の属性例

この要望概念と場所語候補との関連度を算出し、上位 3 件を要望に適した場所語として提案する。表 3 に要望概念と場所語候補の関連度を示す。

表 3: 要望概念と場所語候補の関連度

要望概念	場所語候補	関連度
アトラクション & 乗る	京都	0.230
	志摩スペイン村	0.332
	富士急ハイランド	0.398
	ラスベガス	0.049
	東京ディズニーランド	0.391

表 3 の例では要望概念との関連度が高い「志摩スペイン村」「富士急ハイランド」「東京ディズニーランド」が最終的に提案される。

7 評価

アンケートにより取得した 60 文の要望文をテストセットとして評価を行った。要望文と提案手法により得られる場所語を被験者に提示し、それぞれの場所語について 3 名の目視により判断を行った。入力した要望文に対する提案として適する場所語は○、適さない場所語は×とし、2 名以上が○とした場合を正解とし

た。今回の提案では入力1文に対して3つの場所語を提案するため、テストセット全体で提案される場所語は180語となる。これらのうち、○と判断された場所語の割合を精度とした。図6に評価結果を示す。

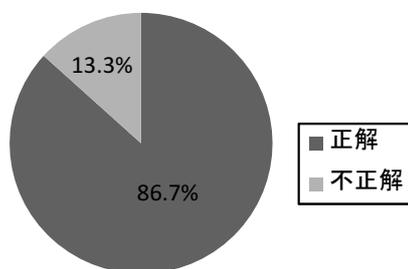


図6: 評価結果

結果として全体の86.7%の場所語が要望文に適する結果となった。また、提案される3つの場所語のうち1つも正解が含まれない要望文は存在しなかった。表4に要望文と提案された場所語の例を示す。

表4: 提案される場所の例

要望文	提案される場所語 (評価)
羊と遊びたい	1:秩父 (○) 2:観光スポット (×) 3:土別 (○)
寺を参拝したい	1:岸渡 (×) 2:神社 (×) 3:京都 (○)
温泉に入りたい	1:汪泉閣 (○) 2:草津 (○) 3:越後湯沢 (○)
九州の温泉をめぐりたい	1:霧島温泉 (○) 2:由布院 (○) 3:杖立 (○)
アトラクションに乗りたい	1:富士急ハイランド (○) 2:東京ディズニーランド (○) 3:志摩スペイン村 (○)

要望文「九州の温泉をめぐりたい」では、提案された3つの場所語が全て具体的な九州の温泉地名となっている。要望文「アトラクションに乗りたい」に関しても場所語は全て具体的な施設名となっており、適した提案がなされていることが分かる。

要望文「羊と遊びたい」から提案された「観光スポット」は広義には場所を表す語であるが、要望から提案される語句としては漠然としすぎているため×の評価

となった。会話のための応答生成へ利用することを考えると、提案されるべき場所語の具体性がある程度必要であることがわかる。

8 まとめ

本稿ではロボットと人間の円滑なコミュニケーションのための技術として、要望文に適合する場所語をWeb上から検索して提案する手法について述べた。Web検索から得られる語句について、シソーラスを用いることで場所語であるかの判別を行った。さらに要望文を構成する語句から要望概念を作成し、場所語候補との関連度を算出することで要望内容に適した場所語を選別した。結果として評価に用いた60文の要望文から提案された180語の場所語について、86.7%の精度を得た。また、すべての要望文について最低でも1つの正しい場所語を提案することが出来た。

謝辞

本研究の一部は、科学研究費補助金(若手研究(B)24700215)の補助を受けて行った。

参考文献

- [1] 奥村紀之, 土屋誠司, 渡部広一, 河岡司. 概念間の関連度計算のための大規模概念ベースの構築. 自然言語処理, Vol.14, No.5, pp.41-64, 2007.
- [2] 渡部 広一 奥村 紀之 河岡 司. 概念の意味属性と共起情報を用いた関連度計算方式. 自然言語処理, Vol.13, No.1, pp.53-74, 2006.
- [3] NTTコミュニケーション科学研究所. 日本語語彙大系. 岩波書店, 1997.
- [4] 辻 泰希 渡部 広一 河岡 司. wwwを用いた概念ベースにない新概念およびその属性獲得手法. 第18回人工知能学会全国大会論文集, 2D1-01, 2003.
- [5] 後藤和人 土屋誠司 渡部広一 河岡司. Webを用いた未知語検索キーワードのシソーラスノードへの割付け手法. 自然言語処理, Vol.15, No.3, pp.91-113, 2008.