

# 現実世界と連動したローグライクゲームの自動生成のための 地理的ナリファレンス情報の Web からの抽出

中山 主税<sup>†</sup> 榮田 基希<sup>†</sup> 荒澤 孔明<sup>†</sup> 服部 峻<sup>††</sup>

<sup>†,††</sup>室蘭工業大学 ウェブ知能時空間研究室 〒050-8585 北海道室蘭市水元町 27-1

E-mail: <sup>†</sup>{13034017,16043009,16043005}@mmm.muroran-it.ac.jp, <sup>††</sup>hattori@csse.muroran-it.ac.jp

あらまし 位置情報ゲームとは、携帯端末の位置登録情報をゲームに利用したものであり、現実世界でモバイルユーザが移動する特徴から観光への応用が注目されている。しかしながら、観光に上手く適用できた成功例は一部に留まっており、これは成功までに課題となる要素が多いからである。そこで本稿では、観光客誘致への応用をより発展させることを目的として、現実世界とより連動したゲーム、特に地理情報との親和性が高いと考えるローグライクゲームの自動生成を試みる。その第一歩として、ゲーム内でアイテム（食糧や回復アイテム）を生成するためには対象地域の名物に関する情報（名前や味、材料など）を、ゲーム内でマップ（部屋や通路）を生成させるためには対象地域の観光名所に関する情報（名前や分類、特徴、評判など）を Web から抽出し、これらの地理的ナリファレンス情報に基づいてアイテムやマップを自動生成するシステムを提案する。また、その抽出精度に関する評価実験を行う。

キーワード 位置情報ゲーム, ゲーム自動生成, 情報抽出, Web マイニング, コンテキスト・アウェアネス

## Geospatial Reference Extraction from the Web for Rogue-like Game Generation Coupled to the Real World

Chikara NAKAYAMA<sup>†</sup>, Motoki EIDA<sup>†</sup>, Komei ARASAWA<sup>†</sup>, and Shun HATTORI<sup>††</sup>

<sup>†,††</sup> Web Intelligence Time-Space (WITS) Laboratory, Muroran Institute of Technology  
27-1 Mizumoto-cho, Muroran, Hokkaido, 050-8585, Japan

E-mail: <sup>†</sup>{13034017,16043009,16043005}@mmm.muroran-it.ac.jp, <sup>††</sup>hattori@csse.muroran-it.ac.jp

**Abstract** A location-based game utilizes geographical location information of a mobile user's device for the game, and its applicability to sightseeing has attracted attention owing to the feature that s/he moves in the real world. However, there are few cases of successfully having applied a location-based game to sightseeing, because it has a bunch of very real problems for success. Therefore, our research tries to automatically generate a game coupled to the real world for the purpose of more successfully applying it to sightseeing, especially a Rogue-like game with high affinity for geographical location information. As the first step, this paper proposes a novel system that mines the Web for information about specialty dishes (their names, tastes, ingredients/materials, and so on) of a target area to generate game items (food and healing items) and also information about sightseeing spots (their names, categories, features, reputations, and so forth) of a target area to generate game maps (rooms and corridors), and automatically generates a Rogue-like game with items and maps based on these geospatial reference information. In addition, this paper performs an evaluation experiment about the precision and recall of geospatial reference extraction from the Web, such as specialty dishes, sightseeing spots of a target area, and their rarity.

**Key words** Location-based Game, Game Generation, Information Extraction, Web Mining, Context-Awareness

### 1. ま え が き

位置情報や地域の情報を利用したゲームは位置情報ゲームと呼ばれ、2000年以降存在するゲームジャンルである。位置情報

を利用することで現実とゲームとの間に繋がりを持たせ新しい体験や視点を提供する、また人が移動することによって可能となる運動不足の解消や人の目が届く場所を広げるといったゲーム外の副次的利用法も模索されている。

位置情報ゲームとの関連で特に注目されるものに観光が挙げられる。ゲーム内での特典やそれに合わせたイベントなどで集客し、それをきっかけに地域の魅力を伝えることを目的としたもので、2015年2月頃に岩手県では位置情報ゲームであるIngress (Niantic, Inc.)を対象に観光PRを行い、多くの参加者を集めた前例がある。しかし、このような試みは一部地域での成功のみに留まっているのが現状である。これは運営する地方自治体における位置情報ゲームへの深い理解が必要となる他、PRと絡める場合の手法の難しさなどの壁が原因と考えられる。この問題の解決を図ることが本研究目的の一つである。

そこで、地域ごとの名物や観光情報を自動取得し、ゲームに反映させることで各自治体のPR企画との関連付けに利用できないか考えた。自動取得することで、具体的な名物を絡めてPRすることが容易になり、技術的知識を持たない人でもPRに利用できる。例としては、地方の名産アイテムが自動生成され、それを収集したプレイヤーは景品を獲得できるといった現実世界とゲーム世界の協調連携が挙げられる。また、プレイヤーに分かりやすく地域の名物を宣伝し、ゲーム内と現実の双方の魅力から集客を見込める等の活用法も考えられる。

本研究では、観光客誘致への利用をより発展させることを目的とする位置情報ゲームの自動生成を試みており、特にログライクゲームに着目している。ログライクゲームは食糧が重要な要素であり、現実世界の観光において重要である名物や土産といった要素と絡めることが容易であり、また、自動生成を核としたゲームであることからリンクする点が多く、地理的位置情報との親和性も高いと考えられる。本稿では、現実世界と連動したログライクゲームの自動生成の第一歩として、ゲーム内でアイテム(食糧や回復アイテム)を生成するためには対象地域の名物に関する情報(名前や味、材料など)を、ゲーム内でマップ(部屋や通路)を生成させるためには対象地域の観光名所に関する情報(名前や分類、特徴、評判など)をWebから抽出し、これらの地理的なりファレンス情報に基づいてアイテムやマップを自動生成するシステムを提案する。

室蘭市を対象に自動生成した場合のゲーム画面例を図1に示す。室蘭市の有名な観光名所として自動取得した「白鳥大橋」を元にゲーム上で海沿いのマップを生成し、これらの観光スポットの地理的位置とゲーム上のマップとをリンクさせる。さ



図1 室蘭市を対象に自動生成されるゲームマップの例

らに、室蘭市の名物として自動取得した「カレーラーメン」を元にゲーム内のアイテムを生成し、これらの名物のPR文をアイテム説明欄に加えることで、現実世界とゲームとを連動させる。

## 2. 提案システム

本章では、最終的な目標システムである「現実世界と連動したログライクゲームの自動生成」の概要、要素や特徴、現実世界との連動に関して述べる。

### 2.1 ログライクゲームの概要

ログライクゲームとは、コンピュータロールプレイングゲームである「Rogue」の特徴を持つゲームジャンルであり、決まった階層までの到達、または限りなく続き、進むごとに難易度が上がる階層をどこまで深く到達できるかを目的とする。決まった階層への到達を目指す場合、最下層にはボスキャラクターが配置されることが多い。「Rogue」の特徴の中で、特に本研究で注目した特徴としてランダム生成されるダンジョン、食糧が重要になるシステムが挙げられる。

通常のロールプレイングゲームでは物語の内容が前提となったマップが作られる。そのため、物語の内容から外れたマップの生成は難しく、対応させる場合は物語の進行自体にまで影響を及ぼしかねない。これに対し、元々ダンジョンがランダム生成で作られるログライクゲームの場合は自由度が高く、様々な地形や道に応じて柔軟に生成の形を対応させることが可能であると考えた。また、観光名所を従来のログライクゲームの形式でダンジョン化させ、用いるデータを観光名所の情報にする場合、また、同じ地域にもう一度来る場合など、同じ場所を複数回遊ぶ可能性が考えられる。この場合、ログライクゲームのランダム生成は同じ場所の周回による飽きを防ぐ。

また、ログライクゲームは、「餓死」の概念が特徴である。ログライクゲーム内では満腹度のステータスが存在し、これは一定行動ごとに減少し、0になった場合は体力が1回の行動毎に減少する。そのまま放置した場合、体力が0になるまで減少し餓死する。これを防ぐために、ログライクゲーム内では食糧としてのアイテムが重要な要素として存在する。この点に着目し、各地域の名物を食糧としてゲーム内で用いることで現実との関連を深める。

### 2.2 現実世界との関連付け

現実世界とゲームを連動させる場合、ゲームデータの要素をリアルに出来るだけ近付ける必要があると考えた。そこで現実世界との繋がりを持たせる為に、まず、ゲームデータの要素である食糧アイテムに地域の名物を用いて、そのアイテムの説明欄に名物のPR文を加える。また、観光名所の地形情報や周辺の特徴からマップを生成する。マップ生成アルゴリズムに関しては、地域の特徴を反映させる際に、危険区域をゲーム内の行動範囲から外して、適切な道を反映させる必要がある。敵キャラクターや味方キャラクターに関しては、マップを生成する際に取得した地形情報や周辺の特徴を元に生成する。また、位置情報を利用することで、現実世界からゲーム側へアプローチをかける。例えば、ゲーム内で登場するアイテムと同じものを提供する飲食店が、近辺に存在するかどうかの情報をユーザに公

開する。実際に現実世界の飲食店などに出向くことで、ゲーム内でのアイテムをクーポンのように使えるようにするだけでなく、飲食店でその地域の名物料理を食べた場合、ゲーム内で同じ料理のアイテムを手に入るようにする。更には、ユーザーキャラクターの操作方法として、手動だけではなく、位置情報を利用した操作方法も考えている。

### 2.3 システム概要

次の図2に提案システムの概要を示す。Step 1 のアプリケーション起動時に、ユーザーがゲーム自動生成を希望する目的地を設定する。次に、Step 2 で、前のステップで設定された目的地を元に、その地域の名物名や観光名所といった地理的リファレンス情報を Web から抽出する。そして、Step 3 で、前のステップで取得された名物名を用いて、ローグライクゲームの重要な要素の食糧アイテムに、地域ごとの名産物や有名な料理をゲームに盛り込む。更に、Step 2 では、その地域の名物名や観光名所だけでなく、その名物の特徴や PR ポイントも Web から取得する。取得された特徴や PR ポイントを用いて、その名物名のアイテムがゲーム内でどのような効果を持つか決定する。また、ユーザーの趣味嗜好によって追加効果を与える。ゲームの中でアイテムとなっている名物や料理を、現実世界でユーザーが実際に購入したり、食べたりのことを SNS やレビュー等に投稿した際にゲーム内で特典が生じるだけでなく、ユーザー間で趣味が似通っている嗜好の人を探す材料と成り得る。その材料を元に、パーソナライズを行い、似た嗜好の人が高評価を付けている名物を優先的に紹介する。また、アイテムの解説欄に PR 文を用いることで、地域への知識を深めると同時に現実世界とゲームにおける繋がりを補強する。

最後に、Step 4 でのマップ生成に関しては、抽出された観光名所を元にマップチップを生成する。例えば、図2のように観光名所が橋だった場合、その地形は海沿いや川といった水辺が考えられる。その特徴を利用して、マップチップの生成は地続きではなく浮き島の地形を用意する。敵の生成に関しても、生成されたマップチップの特徴を参考に敵に持たせる特性を決定する。図2を例にすると、地形は浮き島なので水棲の特性がある敵が用意される。また、味方キャラクターも地域の特徴を元に作成を行い、味方キャラクターのみにおいてカスタマイズ性を持たせる。ダンジョンの攻略や実績解除によってカスタマイズに用いるアイテムを報酬として配布する。「実績解除」とは、何らかのミッション達成やダンジョン攻略といった実績を積むことによって、特別なミッションやダンジョンなどが解放されることである。一部の味方キャラクターには名物のアイテムを与えることにより、ステータス強化等の育成要素を持たせる。この手法は、取得された観光名所が具体的に何を表しているか、また、それが地形とどのような関係を持っているのかを把握するアルゴリズムを考える必要がある。

次に、レアリティ情報の付与に関して説明する。アイテムやマップにレアリティ情報を付与する為に、我々が提案する手法で取得された名物名や観光名所の名詞を用いて、Web 検索を行い、得られた検索結果のヒット数からレアリティを設定する。Web 検索から得られたヒット数が少なかった場合レア度を上

げて、逆に、ヒット数が多かった場合レア度を下げないようにする。このレアリティを設定する手法は、地域ごとにある名物や観光名所の名詞が「目的地の地域の中でどの程度有名であるのか」情報を得る必要がある。

### 2.4 観光記録と別ユーザーへのダンジョン配布

本システムで提案するゲームシステムに、実績やエンブレム等の要素を設ける。実績解除の例として、実際に観光名所に出向くことや、観光名所や名物に対してのレビューや SNS への投稿をした場合が挙げられる。実績解除した場合、ゲーム内で特典を付ける。実際にユーザーが観光名所に出向き観光記録を付けることで、実際に行ったことがある旅行全てをまとめて記録する「旅行記」に近い情報を意図せず作成させる目的がある。「旅行記」を作成することで、まず、「旅行記」に記録されている各観光名所のダンジョンを、自宅好きな時間に何度でも遊べて楽しむことが出来る。ローグライクゲームは、同じダンジョン情報を用いても、自動生成によるランダム性と多様性によって飽きが来ない特徴がある。この特徴から、「旅行記」から日記を読み返すような思い出の振り返りを提供することも目的としている。また、自身の「旅行記」の振り返りダンジョンのみで楽しむだけでなく、ユーザー間で「旅行記」に記録されている振り返りダンジョンの配布を可能にすることで、自分がまだ行ったことがない観光名所のダンジョンのゲームを遊ぶことが出来る。配布されたダンジョンを攻略した場合のユーザーの特典として、配布されたダンジョンを攻略したという実績解除、及び、仲間キャラクターの育成やダンジョン攻略の報酬の収集などが得られる。このダンジョン配布のシステムは、ユーザーから他のユーザーへと観光名所のダンジョンが渡っていくため大規模な宣伝効果が得られると期待される。

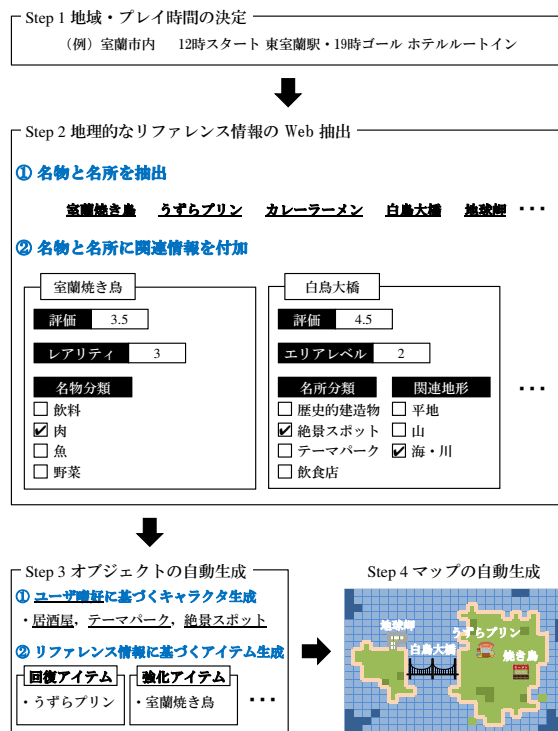


図2 システムの概要

### 3. 提案手法

本章では、ユーザが現実世界と連動したローグライクゲームの自動生成を希望する目的地の名物と観光名所の具体的な名称といった地理的リファレンス情報を Web から取得する手法について述べる。また、抽出された名称からアイテム化、及び、マップチップ化をする際にレアリティを設定する手法も述べる。

#### 3.1 名物と観光名所の名称抽出

ユーザに希望する目的地を設定してもらい、その地名を元に、その地域の名物名や観光名所を Web から抽出する。Web 抽出には Google 検索を用いて、最大 100 件表示させた 1 ページ目の検索結果を利用する。まず、名物や観光名所に関する表現にマッチするパターンを、予め網羅的に用意しておく。そして、予め用意しておいたパターン文に、ユーザが指定した地名を、以下の箇条書きに載っている「検索地名」の部分に挿入して、新たに出来上がったパターン文をそのまま Web 検索に掛ける。

- 名物抽出のパターン文 (40 パターン)
  - 「検索地名」の名物は (名物)
  - 「検索地名」で有名な (名物)
  - 「検索地名」で食べたい (名物)
  - 「検索地名」・・・
- 観光名所抽出のパターン文 (135 パターン)
  - 「検索地名」の観光名所は (観光名所)
  - 「検索地名」のきれいな (観光名所)
  - 「検索地名」で訪れたい (観光名所)
  - 「検索地名」・・・

(名物) や (観光名所) は、パターン文で Web 検索を掛けた時に、パターンに合致した箇所に来ると予測されている、名物名や観光名所の文字列を指している。パターン文の「検索地名」の後に来る形態素には、「名物」「観光名所」などの名詞、「有名」「きれい」などの名詞・形容動詞語幹、「食べる」「訪れる」などの動詞といったパターンが複数個用意されている。Web 検索すると、これらのパターンに合致した Web 文書の検索結果が最大 100 件得られるので、合致した箇所に後続する文字列を形態素解析し、その任意の  $n_m$  個の形態素系列の中から名詞を抽出して取得する。もし名詞が続いて得られた場合は複合語として処理する。その後、「こそあど言葉」などのストップワードは候補から除去する。以上により、検索結果から抽出された(複合)名詞のカウンタを数えて、そのカウンタが予め定めておいた閾値 (以下、 $n_c$ ) 以上の場合、ユーザが設定した目的地の名物名や観光名所として採用する。

#### 3.2 食糧アイテムにおけるレアリティの自動付与

前節の手法で目的地の名産物と判定された名物名にレアリティ情報を自動付与する為の手法について説明する。レアリティ情報を食糧アイテムに付与する為に、[検索地名 AND 名物名] という検索質問で AND 検索して得られるヒット件数を用いる。ヒット件数でレア度を分類する為に、レアリティの閾値を設ける。本稿では、2 つの変数  $t_1$ ,  $t_2$  (但し、 $t_2 < t_1$  とする) をレアリティの閾値として、3 段階のレアリティを設定する。目的地における名物名  $x$  のレアリティ判定を行うための関

数  $r(x)$  を以下に示す。

$$r(x) = \frac{df([\text{場所 AND 名物 } x])}{df([\text{場所}])^p}$$

但し、 $df([q])$  は、検索質問  $q$  で Google 検索を掛けて取得したヒット数である。関数  $r(x)$  はレアリティ判定を行う為の評価値を表しており、 $r(x)$  の数値が高いほど、目的地における名物  $x$  のレアリティを低くする。また、分母の累乗の項  $p$  は単純に割り算を行った場合、目的地に依って分母の影響が強過ぎるため設定したものである。 $p$  の値は、本稿の評価実験では、0.500 から 1.000 の間を 0.001 ずつ変動させて最適化を図っている。最後に、目的地における名物  $x$  の評価値  $r(x)$  に基づくレアリティの判定基準を次の表 1 に示す。

表 1 レアリティ判定の基準

| $r(x)$ に関する不等式        | 目的地における名物 $x$ のレアリティ判定 |
|-----------------------|------------------------|
| $0 \leq r(x) < t_2$   | 3                      |
| $t_2 \leq r(x) < t_1$ | 2                      |
| $t_1 \leq r(x)$       | 1                      |

### 4. 評価実験

本章では、地理的リファレンス情報の抽出手法、及び、名物と観光名所のレアリティの算出手法に関する評価を行う。

#### 4.1 名物と観光名所の抽出手法の評価

本節では、3.1 節で提案された手法によって Web 抽出された地理的リファレンス情報 (名物と観光名所) がどの程度正しく抽出されたかを明らかにしていく。評価には「室蘭」「横浜」「名古屋」の 3 種類の市を用い、第 1 著者でそれぞれの市の名物と観光名所に関する正解セットを予め定めた。例えば、室蘭の観光名所に関する正解には「地球岬」「白鳥大橋」のようなものがある。また、評価尺度には提案手法によって抽出された名詞の適合率、再現率、またこれらの F 値を用いる。表 2 には、3 種類の検索地名の平均名物抽出精度のランキングがパターン毎に示されている。但し、表 2 中の順位は F 値が高かった順に付加されている。また表 3 には、表 2 中の上位 5 件、上位 10 件、及び、全 40 件のパターンを組み合わせた手法による 3 種類の検索地名の平均抽出精度が示されている。表 3 からはパターンを上位 5 件組み合わせる用いた時の F 値が最も高くなり、パターンの数が増加するに伴い F 値が低下している事が読み取れる。本来ならばパターン数を増加させる事で、再現率が上がると期待されていた。しかし、表 2 より F 値が 0.000 となるパターンが全 40 件のうち 8 件存在していた事から分かるように、パターンの検討が不十分であった。これにより、パターンの数の増加に伴い、正解では無いノイズの名詞が多く抽出されてしまい、適合率、再現率ともに低下することになったと考察できる。

表 4 には、3 種類の検索地名の平均観光名所抽出精度のランキングがパターン毎に示されている。また表 5 には、表 4 中の上位 5 件、上位 10 件、上位 15 件、及び、全 135 件のパターンを組み合わせる手法による 3 種類の検索地名の平均抽出精度が示されている。観光名所は名物とは異なり、パターンを上位 10

件組み合わせで用いた時の F 値が最も高くなり、それ以降は名物と同様にパターンが増加するに伴って F 値が低下している事が読み取れる。

次に、提案手法における、パターンに合致した箇所に後続する任意の形態素数  $n_m$  と、切り出された名詞（名物や観光名所の候補）のカウント数に関する閾値  $n_c$  を変動させて、適合率、再現率、F 値への影響を評価していく。図 3、図 4、図 5 には、 $n_m$  と  $n_c$  を変動させた時の 3 種類の検索地名の平均再現率、平均適合率、平均 F 値がそれぞれ示されている。但し、この精度は使用パターンとして表 2 中の上位 5 件のパターンを組み合わせで用いた時のものである。図 3 からは、 $n_m$  と  $n_c$  が底面の黄色い領域の時に最大適合率 0.667 である事が確認できる。また、 $n_m$  と  $n_c$  の値の増加に伴い、適合率が増加している事も読み取れる。図 4 からは、 $n_m \geq 16$ ,  $n_c = 1$  の時に最大再現率 0.289 である事が確認できる。また、 $n_m$  の値の増加、 $n_c$  の値の減少に伴い再現率が増加している事も読み取れる。 $n_c$  の値は、正解以外の名詞の抽出を防ぐ効果を持つ。そのため、 $n_c$  の値を高く設定する事で、確実性が十分である名詞のみ抽出され適合率が増加したと考察できる。しかしながら、 $n_c$  の値を高く設定する事で再現率が低下するという課題も指摘された。

表 2 パターン毎の名物抽出精度比較

| 順位  | パターン           | 適合率   | 再現率   | F 値   |
|-----|----------------|-------|-------|-------|
| 1   | 「検索地名」名物と言えば   | 0.384 | 0.196 | 0.260 |
| 2   | 「検索地名」の名物と言えば  | 0.333 | 0.151 | 0.208 |
| 3   | 「検索地名」の名物は     | 0.433 | 0.135 | 0.206 |
| 4   | 「検索地名」グルメと言えば  | 0.458 | 0.127 | 0.199 |
| 5   | 「検索地名」のグルメと言えば | 0.337 | 0.127 | 0.185 |
| ... | ...            | ...   | ...   | ...   |
| 33  | 最下位の同率 8 件     | 0.000 | 0.000 | 0.000 |

表 3 パターンの組み合わせによる名物抽出精度比較

| パターン            | 適合率   | 再現率   | F 値          |
|-----------------|-------|-------|--------------|
| F 値上位 5 件のパターン  | 0.441 | 0.246 | <b>0.316</b> |
| F 値上位 10 件のパターン | 0.351 | 0.228 | 0.276        |
| 全 40 件のパターン     | 0.363 | 0.202 | 0.260        |

表 4 パターン毎の観光名所抽出精度比較

| 順位  | パターン            | 適合率   | 再現率   | F 値   |
|-----|-----------------|-------|-------|-------|
| 1   | 「地名」の観光名所の      | 0.556 | 0.102 | 0.173 |
| 2   | 「地名」の名所と言えば     | 0.500 | 0.098 | 0.164 |
| 3   | 「地名」のランドマークと言えば | 0.750 | 0.092 | 0.163 |
| 4   | 「地名」の名所         | 0.639 | 0.092 | 0.160 |
| 5   | 「地名」のランドマークである  | 0.287 | 0.109 | 0.158 |
| ... | ...             | ...   | ...   | ...   |
| 63  | 最下位の同率 73 件     | 0.000 | 0.000 | 0.000 |

表 5 パターンの組み合わせによる観光名所抽出精度比較

| パターン            | 適合率   | 再現率   | F 値          |
|-----------------|-------|-------|--------------|
| F 値上位 5 件のパターン  | 0.508 | 0.153 | 0.235        |
| F 値上位 10 件のパターン | 0.444 | 0.207 | <b>0.283</b> |
| F 値上位 15 件のパターン | 0.357 | 0.229 | 0.279        |
| 全 135 件のパターン    | 0.199 | 0.247 | 0.220        |

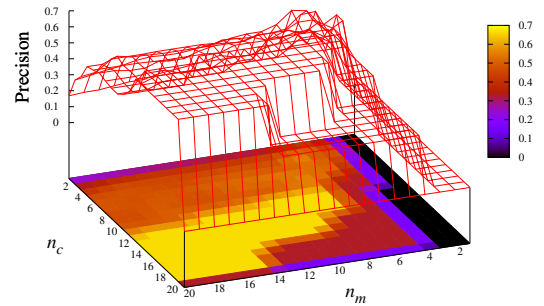


図 3 上位 5 件のパターンを組み合わせた名物抽出手法における適合率

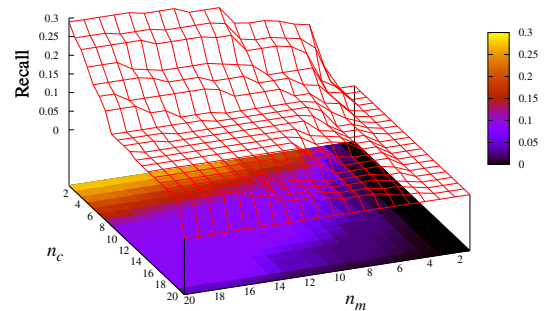


図 4 上位 5 件のパターンを組み合わせた名物抽出手法における再現率

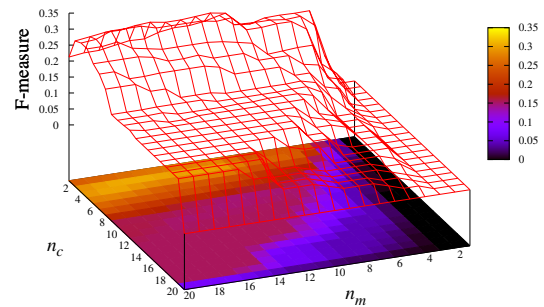


図 5 上位 5 件のパターンを組み合わせた名物抽出手法における F 値

本稿で定めた目的地（検索地名）それぞれの名物には、室蘭市を例にすると、「室蘭焼き鳥」「カレーラーメン」等のようなある程度知られているものは勿論、「草太郎」「げんこつパイ」等のようなあまり知られていないものも含まれている。 $n_c$  の値の増加に伴い、「草太郎」「げんこつパイ」のような名詞に関しては、確実性が十分でないと判断されることが多くなっていき、そのようなマイナーな名物が徐々に網羅されなくなる事から再現率が減少したと考察できる。

#### 4.2 レアリティに関する設定における評価実験

本節では、3.2 節で提案された手法によって算出された名物へのレアリティがどの程度正しく算出されていたを明らかにしていく。この評価実験では、3 種類の検索地名の名物が 3.1 節の提案手法によって正しく抽出されたと仮定しているため、第 1 著者で定めた正解となる名物にレアリティに関する正解も手作業で予め付加した。但し、レアリティは 1 から 3 で表され、

1 から 3 へと大きくなるに連れてレアリティが高くなる。評価尺度には、提案手法によって算出されたレアリティと、予め付加した正解レアリティとの平均二乗誤差 (RMSE) を用いる。

図 6 には、関数  $r(x)$  の分母の累乗の項  $p$  の値を変動させた時の 3 種類の検索地名の平均二乗誤差が示されている。図 6 より、 $p = 0.676$  から  $0.743$  の時に平均二乗誤差が最小値  $0.445$  を取っている事が確認できる。

また、図 7 には、 $p = 0.700$  を固定し、レアリティ判定のための基準値  $t_1$  と  $t_2$  を変動させた時の 3 種類の検索地名の平均二乗誤差が示されている。図 7 より、 $t_1 = 0.13$  から  $0.14$ 、 $t_2 = 0.63$  から  $0.66$  の時に平均二乗誤差が最小値  $0.445$  を取っている事が確認できる。

表 6 上位 5 件のパターンを組み合わせた「室蘭」に対する名物抽出

| 順位 | 抽出された語句   | カウント数 | 適合率   | 再現率   |
|----|-----------|-------|-------|-------|
| 1  | 室蘭やきとり    | 20    | 1.000 | 0.071 |
| 2  | カレーラーメン   | 20    | 1.000 | 0.143 |
| 3  | 焼き鳥       | 11    | 0.667 | 0.143 |
| 4  | やきとり      | 11    | 0.500 | 0.143 |
| 5  | 室蘭カレーラーメン | 8     | 0.600 | 0.214 |
| 6  | 室蘭焼き鳥     | 7     | 0.667 | 0.286 |
| 7  | 豚肉        | 6     | 0.571 | 0.286 |
| 8  | 豚串        | 4     | 0.625 | 0.357 |
| 9  | 草太郎       | 4     | 0.667 | 0.428 |
| 10 | 白鳥大橋      | 4     | 0.600 | 0.428 |
| 11 | 玉ねぎ       | 3     | 0.545 | 0.428 |
| 12 | ねぎ間       | 3     | 0.500 | 0.428 |
| 13 | 洋がら       | 3     | 0.462 | 0.428 |

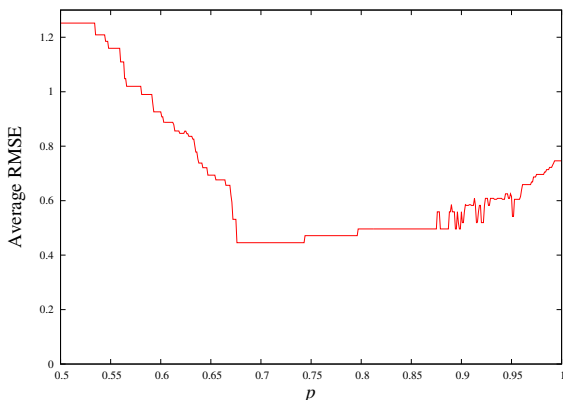


図 6 平均二乗誤差 RMSE と関数  $r(x)$  の累乗の項  $p$

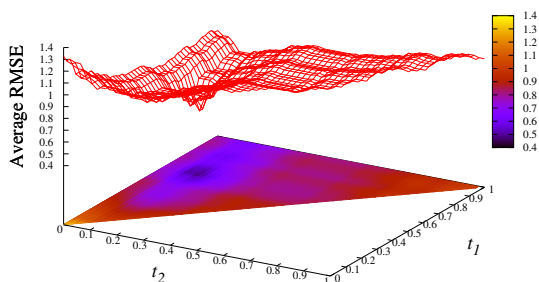


図 7 平均二乗誤差 RMSE とレアリティ判定のための基準値  $t_1, t_2$

## 5. まとめと今後の課題

本稿では、観光客誘致への応用をより発展させることを目的として、現実世界と連動したログライクゲームの自動生成を提案した。その上で、ゲーム内の食糧や回復アイテムを生成するための名物名、また、ゲーム内でマップを生成するために必要な観光名所に関する地理的なりファレンス情報を Web 検索を用いて抽出し、これらの地理的なりファレンス情報に基づいて、アイテムやマップを自動生成するシステムを提案した。評価実験の結果、名物名、観光名所ともに、ゲームアイテムやマップ生成のりファレンス情報として期待した抽出精度は得ることはできなかった。しかしながら、F 値の上位 5 件のパターンを組み合わせることで用いた名物抽出精度や、F 値の上位 10 件のパターンを組み合わせることで用いた観光名所抽出精度においては、比較的優れた抽出精度を実現できる事が確認できた。

今後の課題として、名詞抽出における精度の向上が挙げられる。現在のシステムでは修飾するための助詞「の」で繋がる語句はその都度登録しない限り、正解セットに含まれていた「鐵の素クッキー」や「崎陽軒のシウマイ」のように一つながりの語句として取得することができない。同様に、「ういろう」や「ひつまぶし」のように形態素解析における誤認識の多いひらがなだけの語句に関しても具体例を登録しているため、これらの認識を可能にする必要がある。修飾に反応して単語を処理する方法や、同時出現の回数を記録し、一定の回数や頻度で辞書追加を行う方法を考えている。また、全体的な抽出精度不足が問題である。本稿の実験から、検索語ごとにノイズに関する方向の違い、また用語の表記ゆれが著者の想定より多いことが分かった。これらの反省を元に、全体的な抽出精度不足の解決法を検討する。名物の抽出においては、検索結果に他の地名が現れ、ノイズとなること、またそれら両方の単語を含む検索ワードが存在することを考慮し、それと同時に今後単語に関する要素からアイテム、地域に特徴を持たせる必要性を同時に考える。既に記録している単語の分類を別の検索の際に除外すること、また単語毎の関係を記録するといったことも考えている。

レアリティ判定に関しては、現在の精度では問題がある。本稿では検索地名のヒット数で割り、累乗の項  $p$  によって補正を掛けたが、名物名の総数等は使用しなかった。また、昔から存在するものは知名度が比較的高く、新しく名物として打ち出したものは知名度が低い。これらの手掛かりを利用することで、より細かいレアリティの調整を考えていく。

## 文 献

- [1] 永澤 勇樹, 吉田 京平, 服部 峻, “モバイル端末における旅行記の理解支援のための行程抽出と地図化,” 電子情報通信学会 モバイルネットワークとアプリケーション研究会, 信学技報, Vol.114, No.31, MoNA2014-4, pp.19-24 (2014).
- [2] 中嶋 勇人, 新妻 弘崇, 太田 学, “位置情報付きツイートを利用した観光ルート推薦,” 情報処理学会 データベースシステム研究報告, 2013-DBS-158 (28), pp.1-6 (2013).
- [3] 今田 洋介, 井上 雅史, “GPS ログ収集アプリケーション -ゲームフィクション要素の有用性-,” 第 76 回全国大会講演論文集, 4V-6, pp.3-199-200 (2014).