

行動履歴に基づく探索者の目的推定： 人生のナビゲータ構築に向けて

Goal Estimation of Explorers Based on their Behavior: Toward Development of a Life-navigation System

小林 瞭¹⁾, 日高 昇平¹⁾, 鳥居拓馬¹⁾
Ryo Kobayashi¹⁾, Shohei Hidaka¹⁾, Takuma Torii¹⁾
kaburobo719@jaist.ac.jp

1) 北陸先端科学技術大学院大学
1) Japan Advanced Institute of Science and Technology,

【要約】目的に向けて行動を選択する人に対してその目的達成に有効な選択肢を提示したいと考えた。しかしそのような人が明示的に目的を検索システムのようなものに入力するとは限らず、目的が分からなければ行動の提案もできないという問題がある。本研究では目的に向けた行動を選択する人をマルコフ決定過程モデルから定式化して迷路課題への抽象化を行うことで、行動履歴からその人の目的を推定することで行動選択肢を提示することを考え、具体的に Wikispeedia 探索者の探索過程から行動傾向を探った。このとき人はターゲット到着の前に中間ゴールのようなものを辿ると仮説をたてて検証を行い、結果として仮説を支持する要素を発見し、目的状態の判別への有効性が確認できた。

【キーワード】目的推定, 意思決定, マルコフ決定過程, 限定合理性, 意味知識

1. はじめに

我々は何か目的を決めたとき、その目的を達成する行動選択肢の中でも最良の行動を選択しようと考ええる。しかし、このときの行動選択肢というのは我々が思いつく選択肢に限られてしまうため、実際には存在するが気づいていない選択肢を含めた全ての選択肢の中で最良の行動を選択するのは難しい。さらに、一度自分の中で目的達成に最良だと思い込んだ行動選択肢が存在すると、他の行動選択肢を検討する必要が無いように思えるため、全ての選択肢の中から最良の行動を知る機会を失ってしまう。つまり、自分自身で思い描いた行動選択肢よりも良い行動が存在するかもしれないのに、自身のもつ行動選択肢が最良のものだと思い込んでしまうことが多いのである。そこで、自分でより良い行動選択肢に気づきにくいのであれば、自分よりも目的達成に良い行動選択肢を知る人から最良の行動を提示してもらいたいと考えた。

行動を提示するときに同じ行動でも目的によって良し悪しが分かれるように、行動の良さの基準は目的の内容に依存するため、人の目的を知る必要がある。このときに人が検索等のシステム上で目的の明示的な入力があれば良いが、自身の持つ行動選択肢が最良だと思い込んでしまうため、その入力自体が行われないと予想される。そのため人の目的を推定することから行動選択肢の提示を行うことを考えていく必要があると考えられる。以上より本研究の最終的に目指すことは、目的に向けて行動する人の目的推定を行い目的達成に有効な選択肢を提示するシステムを構築することである。

本研究ではシステム構築までの第一歩としてまず、目的に向けて行動する人の行動履歴から、目的推定を行うことを研究目的とする。目的に向けて行動する人の行動をマルコフ決定過程として定式化する。その定式化の下で、人の十分ではない知識に基づく探索行動を迷路課題とみなし、迷路課題での目的推定課題について検討する。次に、実際に目的に向けて行動する人の行動データとして Wikispeedia データを用い、人の行動原理として「人は最終的な目的の前に、その目的につながる中間ゴールを経由地として目指す」という仮説の検討を行う。

本稿では、まず2節で目的に向けて行動する人の行動についてマルコフ決定過程モデルを用いた定式化を説明し、迷路課題への抽象化したときの目的推定課題について触れる。3節では実際に人の行動データとして用いた Wikispeedia データについて説明する。4節では本研究で目的に向けて行動を選択する人の行動仮説について述べた後その仮説を示す要素を Wikispeedia データの分析を行うことで検討する。5節では4節で検討した要素を用いて Wikispeedia を行う人の行動から目的推定を行う。また、ターゲット記事絞り込みを行い今後の仮説を立てる。最後に5節で本研究の結論と今後の課題について述べる。

2. 目的に向けた行動の捉え方

2.1 マルコフ決定過程

本研究では人の行動履歴は、行動選択肢が状況に依存し、また状況はその前に取った行動に依存する、という行動・状況・行動選択肢の連鎖的依存関係を仮定する。仮定を定式化するためにマルコフ決定過程（Markov Decision Process : MDP）モデルを用いる。MDP モデルは、エージェントがある状態から行動をして次の状態へ遷移し、同時に遷移後の状態から報酬を得るというエージェントと状態の相互関係として定式化している。図 1 は家にいる人がリラックスするという目的を達成するまでに潜在的にとりうる行動を状態遷移図を用いて表す。図 1 では状態「家にいる」から行動「電車に乗る」を選択し次の状態「電車に乗った」に移る。そしてその状態から選択できる行動「歩く」「バスに乗る」が与えられ、選択を迫られる。そして最終的に状態「カフェに到着」から「リラックスする」を確定的に達成している。このように人の行動について、ある状態から行動を選択し次の状態から行動し、さらにその行動が状態を変えるとといった連鎖構造として表せる。1 節で述べた本研究で目指すシステムについて図 1 を用いて説明すると、暗黙的にリラックスするならカフェへ行くと思い、明示的に「リラックスする」という目的が与えられないまでもカフェへ行く人を対象にしている。これは図 1 では状態「カフェに到着」と状態「リラックスする」までの行動が暗黙的になっているものとして考えられる。そして、このような人に対して、暗黙的となっている「リラックスする」を推定し、「リラックスする」を達成するならカフェ以外にも映画に行くことなどをお勧めしてほしいと考えている。

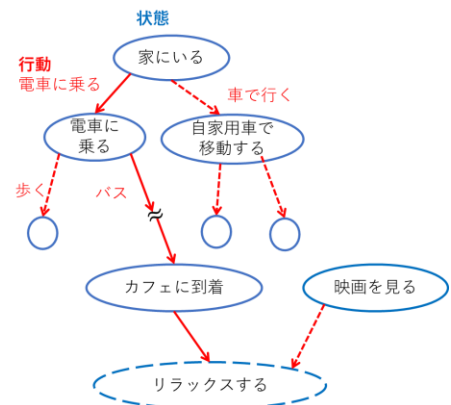


図 1：マルコフ決定過程を用いた表現

2.2 迷路課題への抽象化

迷路課題とは探索者がスタートマスからターゲットマスを目指してマスを移動する課題である。各マスには報酬値が設定されており、この報酬値をできるだけ多く取得することが探索者に求められている。探索者は進むマスを決める方策をもつ。図 2 は迷路課題を分かりやすく二次元格子状迷路で示したものであり、探索者 A がスタートマスからターゲットマスまでマスを移動した後について表しており、赤点線矢印は探索者 A の行動履歴である。マスの色は探索者 A から迷路を見たときにスタートマスとターゲットマス以外は報酬値が分からないことを表している。マルコフ決定過程との対応関係も表現が可能であり、例えば図 1 で最初の状態「家にいる」は図 2 ではスタートマスとして対応でき、状態「リラックスする」はターゲットマスとして対応できる。また図 1 で矢印を選択し行動することは図 2 の迷路課題では方策に従ってマスを移動することである。

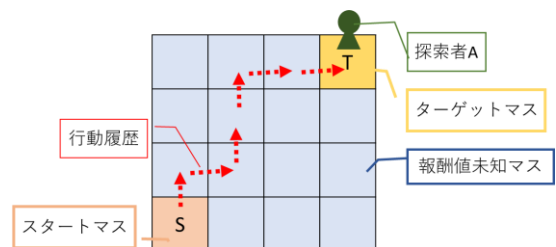


図 2：迷路課題の探索者

本研究では人に対して行動選択肢の提示を行うためにその人の目的を推定する必要がある。具体的に推定にはその人の行動履歴を参考にできることを 1 節で述べた。人の行動を迷路課題に抽象化すると探索者の行動履歴からターゲットマスの推定を行うことが研究目的の達成に繋がる。このとき、探索者に行動の提示を行う人のことを案内人とする。案内人に与えられる情報は探索者の行動履歴と迷路構造のみである。迷路構造とはマスの繋がりのことをここでは指す。現実の人を模した探索者はターゲットマスまでランダムに進むことはなく、不完全ながらも迷路の知識をもって行動方針として方策を決め、迷路課題を解いていると考えられる。そのため本研究ではこの方策を知ることで探索者のターゲットマスを推定できると考えている。これらの課題に対して強化学習(Sutton & Barto, 2000; Babes, Marivate, Subramenian & Littman, 2011)や逆強化学習(Ng & Russell, 2000; Abbeel & Ng, 2004; 五味, 2017; Ho, Littman et al., 2016)を用いることで人の探索過程や、人の行動からその人の価値判断について分析することがで

きる。しかし、本研究では人の行動選択にはその人の思い込みがある事に注目して、その思い込んでい
る選択肢以外に良い選択肢があればそれを提案することを考えていた。そのため実際に人の行動データ
として迷路課題を用いることは実験設計上難しい。そこで、続く3節では迷路課題の特徴を持ちつつ人
の思い込みについても実験条件として満たしている Wikipedia について述べる。

3. Wikipedia：不確実な知識を基にグラフ上を探索する課題

3.1 Wikipedia データを利用する理由

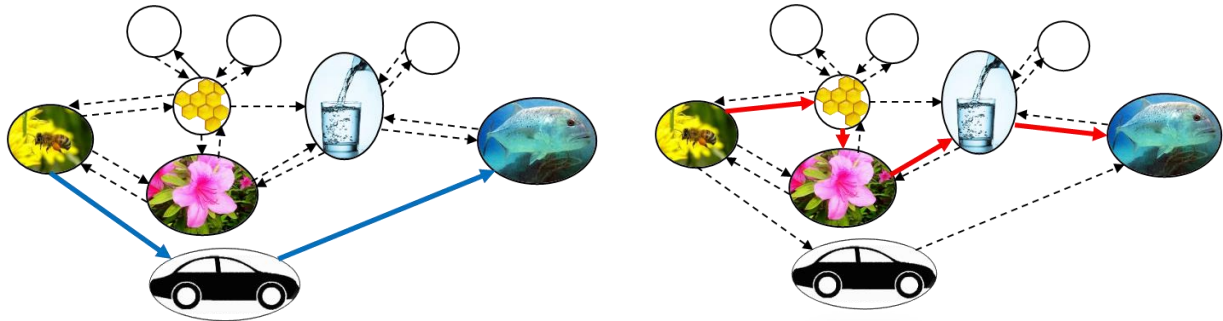


図 3：Wikipedia 記事リンクの例

画像は左上から Bee, Honey, Flower, Car, Water, Fish である

左：Wikipedia リンク構造と最短経路（青太矢印） 右：探索者の行動履歴経路（赤太矢印）

本研究では迷路課題と同様の性質を持ち、かつ人のその思い込みなどのため必ずしも最適な行動をとっているとは限らない探索行動のデータセットとして Wikipedia データ(West, Robert, Pineau & Precup, 2009)を利用する。Wikipedia 上の人の探索行動データを分析することで、目的に向かう探索者の認知過程を調べる。まず Wikipedia というゲームは、ある一回の試行で探索者に Wikipedia 記事ページのある二つがスタート記事とターゲット記事として与えられ、スタート記事から記事中のハイパーリンクのみを辿って、ターゲット記事まで到着することが課題である。そして、スタート記事からハイパーリンクを辿って別の記事に遷移することを1ステップと呼ぶとき、このステップ数をできるだけ小さくすることが探索者の目標として与えられている。このとき、探索者にとってスタート記事とターゲット記事を最短経路でたどることが最適解となるが、現実には最短経路をたどる場合は少数であり、ほとんどの場合冗長な経路をとる。なぜなら、記事リンクをある程度人は想像できるが、いくつかのリンクについては人には繋がっていると想像できないものが存在するからである。例えば図3は Wikipedia 記事のうちスタート記事 Bee とターゲット記事 Fish までのリンク構造の一部を示している。丸で囲まれた絵は Wikipedia 記事を示し、破線の矢印は記事から記事へのリンクの向きを示している。図3の左図では Bee から Fish へ向かうときに Car を経由する青の実線が最短経路となることを表している。これはもし Wikipedia 記事のリンク構造を把握できていれば最短経路を選択することが可能となる。一方で図3の右図は実際に探索者が通りやすい経路を示している。左図で示された最短経路であるはずの青を経由しないことは、探索者にとって Bee から Car へのリンクはあるが Car が Fish までの最短経路とは思わなかったためにその経路を辿っていないことを表している。そのため赤実線矢印のように Honey, Flower, Water を経由することが多い。このように人が事前に想定するリンク構造と実際の Wikipedia のリンク構造には乖離がある場合がある。従って、Wikipedia のデータは、課題に関する不完全な知識(限定合理性)をもつ探索者が目的へと向かう行動履歴を反映したものにとらえることができる。以上より、Wikipedia データを利用する目的は、探索者が不完全な知識しか持たない中で、探索者にとって最適解に近い経路をとろうとするときの行動原理を探ることである。

3.2 暗黙的な目的と明示的な目的の関係

Wikispeedia では探索者はターゲット記事を知っているため 1 節で述べていたような「リラックス」するという目的が明示的に与えられているものとして考えられる。このような場合でも、おそらく人はその目的に直接向かうのではなく、それにつながると思われる代表的な経由地(中間ゴール)を想定しながら、行動するのではないかと我々は仮説を立てた。例えば、図 3 のケースでは、Fish に意味的に近く、かつ Bee などの他の多くの記事とつながっている経由地として Water を想定し、Water まで向かうことができれば実質的に Fish にたどり着くのは容易であると計画を立てることに相当する。これは、特定の人にとって、「カフェ」(代表的な経由地)まで行けば「リラックス」(本来の目的地)できるに違いない(図 1)と考え(場合によってはこれは思い込みである)、「リラックス」の代理として「カフェ」に向かうのと同等である。

従って、明示的な目的を持った探索行動の中に「経由地を想定した探索行動」が現れるとすれば、それを繰り返すうちに、経由地を目的地の代わりに目指す(目的の暗黙化)という学習が行われると考えられる。そこで、明示的に目的を持つ探索行動データとしての Wikispeedia 分析を行い、暗黙化される経由地を目指す行動の発生を念頭に分析を行った。

3.3 Wikispeedia データ構造

研究では目的推定法について検討するために Wikispeedia データを利用する。この節では使用する Wikispeedia データの内容や構造について説明する。まず Wikispeedia で使用した記事の総数は 4,604 記事存在し、データ全体のスタート記事とターゲット記事はこの 4,604 記事の中からランダムに決定される。記事間のリンクとはハイパーリンクで繋がれた記事を指し、全体で 119,882 リンクがある。例えば図 3 のように記事 Bee は Honey や Flower へのハイパーリンクを持っている。このとき全ての記事で相互にハイパーリンクが存在するとは限らないことに注意する。図 3 の場合 Honey から Water へリンクがあるが、Water から Honey へのハイパーリンクは存在しない(矢印の元の記事から矢印の先の記事へハイパーリンクが繋がっている)。カテゴリ構造とは各記事に与えられた抽象度の階層である。試行数とは、ある探索者にスタート記事が与えられてから、ターゲット記事にたどり着くかリタイアするまでの一連の行動を一つのエピソードとしたときのエピソード数である。試行数のうち成功した数は 51,308 回、リタイアか時間制限によりターゲット記事までたどり着かなかった数は 24,875 回であった。ある 2 つの記事の間の最小遷移ステップ(最短経路長)は $4,604 \times 4,604$ の行列で表現できる。全ての記事間のステップ長さは 9 回以下でありデータのうち全体の 98% が 2 回から 5 回ステップでターゲット記事までたどり着いた。強連結グラフとは、どの 2 つの記事の間にも両方向に経路が存在する(部分)グラフを指す。この分析では、Wikispecia の最大強連結部分グラフに含まれる記事のみ対象とし、つまりハイパーリンクで遷移できない記事を持つ記事を除外している。強連結グラフに含まれる記事に注目する理由は、人が記事遷移を行う様子が強連結している記事の数によってチャンスレベルなどの変化により一概に分析することが難しいと判断したためである。強連結記事群は総記事 4,051 記事、そのうちスタート記事は 4,014 記事、ターゲット記事は 3,325 記事、強連結記事の一つの記事にあるハイパーリンク数の平均は 27.6 であった。

4. 目的に向けた行動を行う人の行動モデル

人は不完全な知識をもとに行動しているため、目的まで最適に行動を選択できるとは限らない。この場合の人の行動傾向を定式化するため、本研究では人は目的に向けた行動として中間ゴールのようなものを經由していると仮定する。中間ゴールという考えは車で目的地に向けて道を選択する様子からアイデアを出した。例えば車で北陸先端科学技術大学院大学(JAIST)から金沢駅近辺の店に行くとき、詳細な知識をもたない観光客などはバイパスのような大きな道を通り、金沢中心部を中間ゴールとして目指し、金沢中心部に至った後に目的地周辺の細い道を選ぶようになるだろう。その地域に関する土地勘を持っている者でない限り、最初から住宅街のような細い道を進むことは少ない。最初からではなく最後に細い道を選ぶ主な理由は、ほぼ確実に店の近辺に着けると考えられるまでできるだけ道に迷いたくないからである。最初から住宅街を通っていた場合、道を少し間違えてしまうとそこから修正するのに時間がかかるが、最初に金沢へ通じている太い道を進んでいけば大きく道を外れない限り目的地の近辺までへ向かえる。このように、中間ゴールとは目的にほぼ確実に向かえる局所的な経路の入り口にあたる経由地を指す。図 3 の例では Bee から Fish まで遷移するとき、一見何の繋がりを持たない 2 つの記事を繋げ

るために、Water まで行ければ Fish までほぼ確実に向かえると考えた場合、Water が中間ゴールとなる。

以降では中間ゴールを経由するという仮説を Wikispeedia データの分析から検証する。以降で取り上げる要素については West ら(2012)の先行研究を参考に検討を行う。なお、本節でのデータ分析は West らの追実験である。先行研究で West らはウェブ上を知りたい情報を求めてネットサーフィンする人の行動を、Wikipedia 上をハイパーリンクで動く人の行動として抽象化し、目的とする Wikipedia 記事の推定を行うことで、間接的に人のウェブ上での目的推定を行うことを考えていた。そして West らは Wikipedia 記事とその遷移に表れる要素の特徴ベクトルを生成しターゲット記事推定を行った。結果として West らの行ったターゲット記事推定では、Term Frequency – Inverse Document Frequency (TF-IDF)、Degree、Shortest Path Length の 3 つを特徴ベクトルの主要要素であるとしている。TF-IDF とは文書内に出現する単語の重要度を表す。Degree とは、ある記事に他の記事からリンクの数(次数という)を指す。他の記事からある記事へのリンク数を Indegree、ある記事から他の記事へのリンク数を Outdegree という。Shortest Path Length (以下 SPL) とは Wikipedia の記事間を繋げるリンクを辿ったときの最小遷移数を表す。以上の要素に加えて本研究では West ら(2009)が取り上げた PageRank も検討する。PageRank とは記事遷移を存在する選択肢を一様確率で遷移したと想定したときのある記事に滞在する定常確率のことを表す。以上より本研究ではこれらの要素の中で Degree と SPL、PageRank を中間ゴールの特徴を表す要素として用いることを検討する。ただし、本研究では現実における人の行動を抽象化したものとして Wikispeedia データを用いているため、我々の行動が文章で表されるとは限らないことから TF-IDF 分析は行わない。

4.1 最短経路長減少傾向の分析

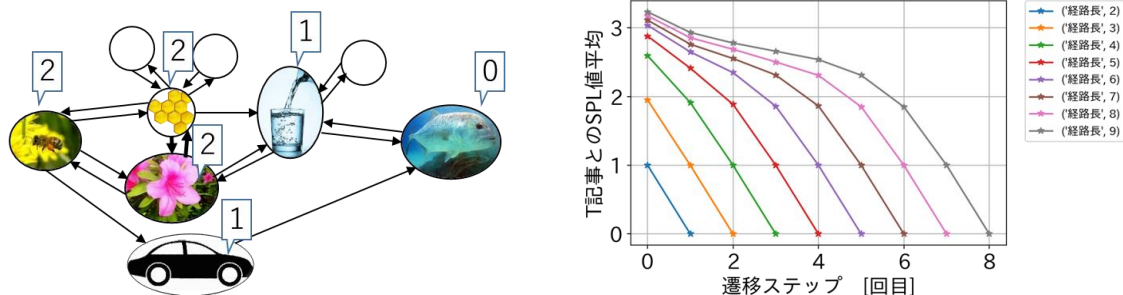


図 4：最短距離減少傾向の分析

左：Wikispeedia リンク構造と最短距離（吹出は Fish までの最短距離）

右：各経路長での探索者の遷移した記事からターゲット記事までの最短距離平均

Shortest Path Length (SPL ; 最短経路長) とはターゲット記事まで最も少ないステップ数での遷移回数を表す。図 4 左図は最短距離減少傾向の分析のイメージ図と結果を表し、具体的に左図はスタート記事 Bee からターゲット記事 Fish としたときに各記事と Fish までの最短リンク数を吹き出しで示している。Bee は Honey を経由した場合の距離は 3 となるが、Car を選択できれば距離 2 となるため、吹き出しには 2 が書かれている。本研究では中間ゴールとは、そこを経由後に最終的なゴールまでは最適に近い経路をとることができるような経由地である。したがって、もしそのような中間ゴールがあるならば、一定のステップ以降には、ページ遷移のたびに最短経路長を縮めるような結果が見られるが、それ以前では最短経路長を縮める速度が遅い、ということが予測される。従って、以下のような予測が成り立つ

予測 1 (中間ゴールあり): ターゲット記事の数ステップ前から最短経路数をより下げる遷移が見られる。

予測 2 (中間ゴールなし): 最短経路をどのステップでも一定程度小さくするような遷移が見られる。

図 4 右図は、各経路長での探索者の遷移した記事からターゲット記事までの最短距離平均を示している。経路長とは人が一試行でスタート記事からターゲット記事まで遷移した記事数である。横軸はスタート記事を 0 ステップ目と数えて 1 遷移するごとに数値が 1 増える。

結果：経路長が長くなるほどスタート記事に近い記事遷移では SPL が 2.0~3.0 の記事を少しずつ減少するよう遷移していることから、ターゲット記事まで最短経路で一直線に進んでいないことが分かる。一方ターゲット記事直前の 2 個前ほどから SPL が急に下がっていることが分かる。

考察：結果よりターゲット記事まで一定に最短経路を選択しておらず，一方でターゲット記事の2ステップ前から最短経路を進めていることから，中間ゴールがある場合の予測1の方が，中間ゴールがない場合の予測2よりも結果に近いことが分かった．よって予測1からは中間ゴールまでは探索するように人は記事遷移を行い，中間ゴールの後は最適な行動を選択できていることが示唆された．

4.2 Outdegree 傾向の分析

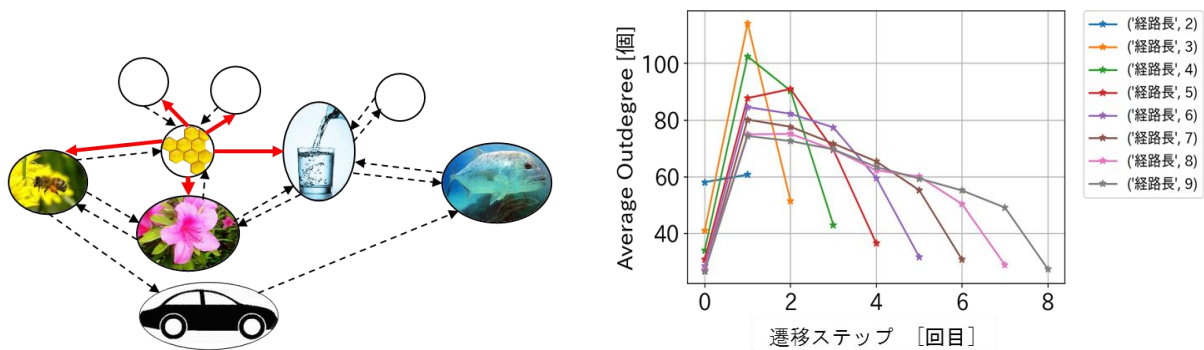


図 5：Outdegree 傾向の分析

左：Wikipedia リンク構造と記事「Honey」の Outdegree の例(赤実線矢印)
右：各経路長での探索者の遷移した記事の Outdegree 数

Outdegree とはある記事から他の記事へのリンク数を指し，ある状態からその人が選択できる行動選択肢の数を表す指標の一つである．図 5 は Outdegree のイメージ図と分析結果を表している．図 5 の左図は一例として記事 Honey の Outdegree を赤色実線の矢印で示している．図 5 右図は，各経路長での探索者の遷移した記事の Outdegree 数の平均を示している．横軸はスタート記事を 0 ステップ目と数えて 1 遷移するごとに数値が 1 増える．

もし中間ゴール前よりも後で潜在的に迷う(最短経路でない記事を選択する)可能性が小さくなるならば，Outdegree はターゲット記事の一定ステップ前に減少する傾向があると考えられる．

予測：中間ゴールの後は，選択肢(Outdegree)が減少する．

結果：Outdegree が多い記事へ遷移した後，ターゲット記事の 2 ステップ前の記事までその数を緩やかに減らし，最後の 2 つの記事では Outdegree が急速に減少することが分かる．

考察：結果より選択肢の数をターゲット記事まで緩やかに下げるように遷移し，ターゲット記事の 1 つか 2 つ前で急に下がることから，中間ゴールを経由することを仮説とした予測を示す傾向が見られた．一方で，探索者がターゲット記事直前の記事でほぼ最適な行動を選択できる理由として，選択肢個数の減少が影響しているのか，それとも探索者の持つ単語繋りの知識が影響しているのかを分けることはこの結果のみから判断することは難しいと考えられる．

4.3 PageRank 傾向の分析

PageRank(Page, 1999)とはスタート記事やターゲット記事関係なくその記事を通る確率を指す．Outdegree が一回の遷移について見ていたのに対して，PageRank はある記事から他の記事までの複数回の遷移について見ている．これはある状態から遷移できる他の状態，更にその次に遷移できる状態への経路の多さの指標の一つとして扱える．図 6 は PageRank のイメージ図と分析結果を表している．図 6 の左図は記事 Flower と Water が PageRank の高い記事として太枠線で示されている．PageRank の考え方として，例えば図 6 の左図では Honey よりも PageRank の高い Flower は，Fish へリンクする経路を Honey よりも多く持つ可能性があると考えられることができる．ここで，実際には Honeyの方が Flower よりも Fish へ向かう選択肢が多いこともあり得ることであり，Flower が Fishに限らず他の記事との経路を多く持つことを表していることに注意する．図 6 右図は，各経路長での探索者の遷移した記事の PageRank の平均を示している．横軸はスタート記事を 0 ステップ目と数えて 1 遷移するごとに数値が 1 増える．

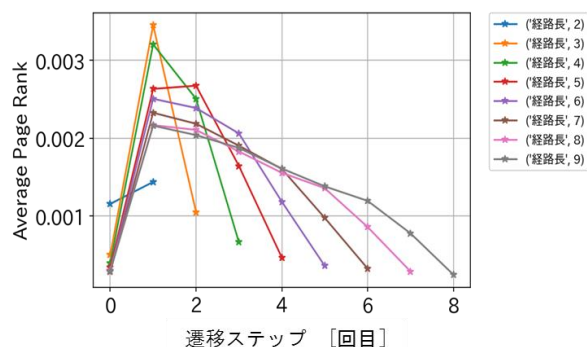
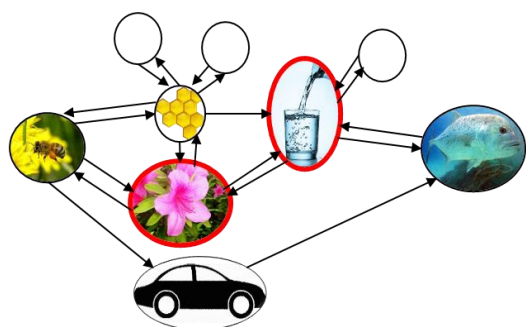


図 6 : PageRank 傾向の分析

左 : Wikispeedia リンク構造と PageRank の高い記事(赤太線)

右 : 各経路長での探索者の遷移した記事の PageRank

予測 : ターゲット記事直前の数ステップ, より PageRank の少ない局所的な記事を選ぶ.

結果 : PageRank をターゲット記事の 2 ステップ前まで維持するように遷移し, 2 ステップ前からターゲット記事までは比較的 PageRank が低い記事を遷移していることが分かる.

考察 : 結果よりスタート記事から最初に PageRank の高い記事へ遷移することと, PageRank を比較的下げるターゲット記事 2 ステップ間が図 3 での考察のように, ターゲット記事まで最適で遷移できていると考えられることから予測と一致する結果と言える. 人がまずは中間ゴールへ向かうために, 現在状態よりも多くの状態へ繋がっている可能性のある状態へ向かう様子を示していると考えられる. ただし, 多くの状態へ繋がっていようとそこにターゲット記事までの経路が含まれていなければ, 最短ステップ遷移になるとは限らない.

5. 目的推定課題

中間ゴールを経由しているという行動傾向に相関があることが示唆された SPL と PageRank を説明変数としターゲット記事予測を行う. Outdegree は PageRank との相関係数が 0.858, p 値が 0.000 であることから PageRank とほぼ同じ傾向を表す指標として考え, 回帰係数の個数が冗長になることを防ぐために説明変数に加えない. ターゲット記事予測のために次の順番で推定が必要だと考えた. まず, 探索者がターゲット記事に既に到着していることを予測する. なぜなら既にターゲット記事に到着している探索者に対してターゲット予測を行う必要は無いからである. 次にまだターゲット記事へたどり着いていない場合に対して, 複数記事をターゲット記事として絞り込むように推定を行い, 絞り込みの精度について検討する. さらにそれぞれの推定に回帰分析を用いて, 推定に用いた説明変数の正答率への寄与度についても検討する.

5.1 終端記事判別

終端記事判別とは, Wikispeedia データのある一試行を抜き出したときに探索者がいる記事が探索者にとってターゲット記事なのかどうかを判断することである. 図 7 は終端記事判別に用いるデータについて説明する. スタート記事 (S) を左端, ターゲット記事 (T) を右端として左から右に向かって探索者が遷移する. 問題設定のために使用するデータは, ターゲット記事に到着済みの試行 (図 7 上の行) と, その試行と同数の遷移を行ったターゲット記事に未到着の試行 (図 7 真ん中の行) である. 後で述べる学習データを示す「窓」に含まれる記事のうち, 最後に遷移を行った先の記事を終端記事 (図 7 青枠の右端記事) と呼ぶ. 学習に用いるデータは終端記事から数ステップ遷移スタート記事へさかのぼるよう選択する. 使用するデータは, 1 試行の中でスタート記事からターゲット記事までの総記事数と, 窓に含まれる記事数を変数として作成する. 例えばスタート記事からターゲット記事までの総記事数が 6, 窓に含まれる記事数を 3 としたときは, 総記事数 6 と 7 の試行の内終端記事から 3 回さかのぼるまでに含まれる記事についてデータを作成する. これらの学習データを回帰係数としてロジスティック回帰モデルを用いた分析を行う. 学習を行うシステムに与えられる学習データの Input は, 図 7 の実線青枠の窓に含まれる記事に対して SPL 比較と PageRank 値を計算したものであり, 教師データとして最後に訪れている記事がターゲット記事であるかどうか与えられる. このとき, SPL 比較では $SPL(u_i, u_T) \geq SPL(u_{i+1}, u_T)$ から True で 1, False で 0 を返す. このとき u_i とは整数 $i (i \geq 0)$ 番目の記事を指し u_T はターゲット記事を指す. $SPL(a, b)$ は記事 a と記事 b の SPL 値を表す. そしてテストデータ (図 7 下の行, 点線青枠窓) には窓に含まれる記事と同じ数の記事について与えられ, テストデータの最後の記事がターゲット記事であることを判断した結果を Yes か No (1 か 0) で Output する. テストデー

タには終端記事がターゲット記事であるものとそうでないものを同数含めることでチャンスレベルを 0.5 にしている。

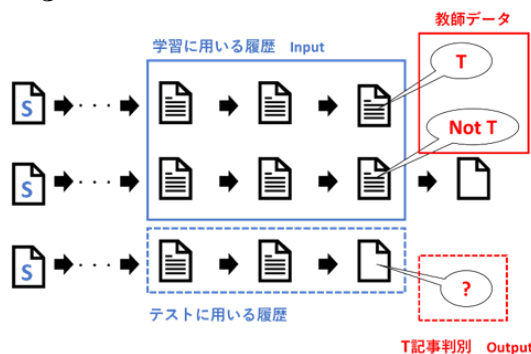


図 7：終端記事判別に用いる学習データ

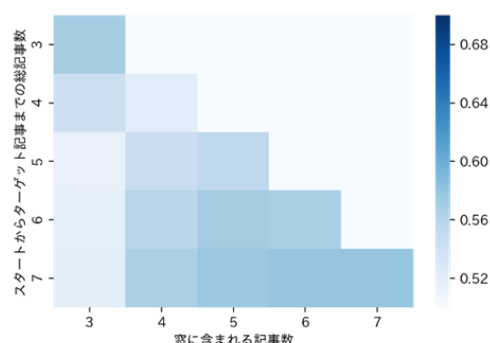


図 8：終端記事判別正答率

結果として図 8 に終端記事がターゲット記事を判断したときの正答率を示す．横軸が窓に含まれる記事数，縦軸がスタート記事からターゲット記事までの総記事数である．図 8 右上部分は窓に含む記事数が総記事数以下となるため計算は行われていない．図 8 より窓に含まれる記事数が大きくなるほど正答率が上昇することが分かる．一方で窓に含まれる記事数が 3 のとき，総記事数が 3～7 の場合に正答率が下がった．このことから総記事数が大きくなるほど終端記事から 3 ステップさかのぼる記事まで見たデータではターゲット記事判別が難しいことが分かる．また，窓に含まれる記事数が 4 を超えると総記事数が大きくなっても正答率が上がることから，終端記事から 4 ステップさかのぼる記事まで見たデータにターゲット記事判別に有効な要素が含まれていると考えられる．

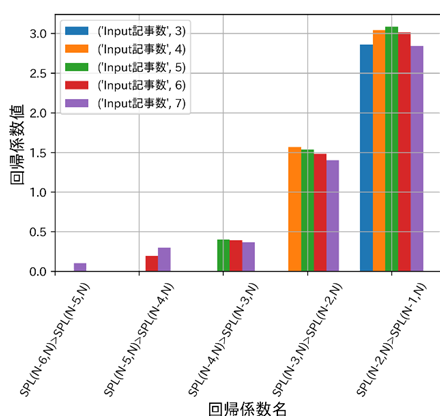


図 9：終端記事判別に用いた回帰係数 SPL 比較

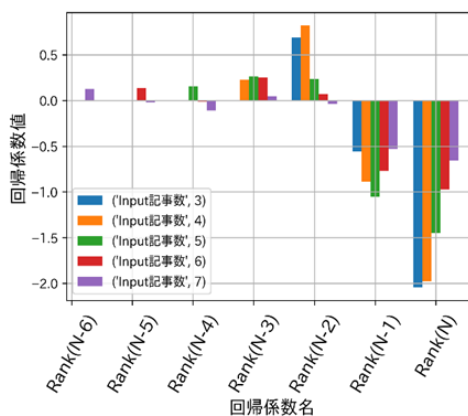


図 10：終端記事判別に用いた回帰係数 PageRank

次に，判別を行うために用いた回帰分析の係数を比較する．図 9 は窓に含まれる記事数ごとに SPL を比較した係数の重みの比較について表している．図 10 は窓に含まれる記事数ごとに PageRank の係数の重みの比較について表している．終端記事を N 番目としたとき，N-1 番目の記事は終端記事からスタート記事へ 1 ステップさかのぼった記事を指す．図 9 から終端記事と 1 つ前，2 つ前記事との SPL 比較の重みが大きくなっていることが分かる．これは図 4 で示されているように 中間ゴールは SPL の急な減少の開始点であることがターゲット記事判別に寄与していることが分かる．また，図 10 より終端記事 1 つ前と 2 つ前の PageRank に重みが付いていることから，中間ゴールを過ぎた後の遷移記事の PageRank にターゲット記事に近いことを示す意味が含まれていることが示唆された．

5.2 ターゲット記事絞り込み

ターゲット記事絞り込みとはある試行の記事遷移履歴を用いて、Wikispeedia に含まれる全ての記事からターゲット記事を絞り込むことである。ここでは 4.1 節で述べた SPL 減少傾向から、探索者の遷移した記事のうち、SPL が大きくなった記事をターゲット記事候補から除外することで、ターゲット記事絞り込みを行った。

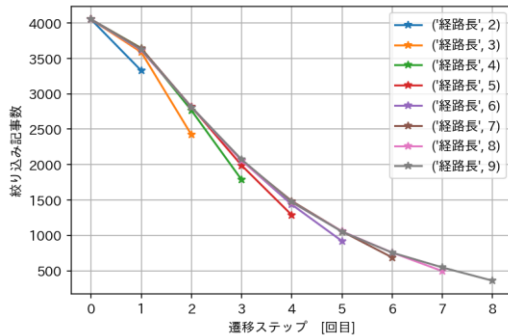


図 11：ターゲット記事絞り込み記事数

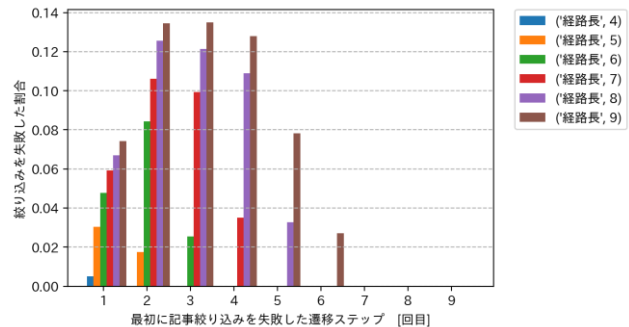


図 12：絞り込みを初めて失敗した遷移ステップ

結果として図 11 のように各経路長で記事の絞り込みが行われ、絞り込みに成功している確率が経路長で異なった。2 回, 3 回経路長では全ての試行で絞り込みが成功し、それ以降は経路長 4 で 0.99, 経路長 5 で 0.95, 経路長 6 で 0.84, 経路長 7 で 0.70, 経路長 8 で 0.54, 経路長 9 で 0.42 と徐々に絞り込み成功率が低下していた。これより記事の絞り込みは必ずしも成功するとは限らないと分かった。そして経路長が長くなるほど絞り込みを失敗していることが分かる。次に、記事の絞り込みを失敗してしまったステップについて分析を行い、結果を図 12 に示す。図 12 はターゲット記事絞り込み失敗率を縦軸、ある試行の中で最初に記事の絞り込みを失敗した遷移ステップ数を横軸として分析している。図 12 から中間ゴールの手前に相当する遷移部分で絞り込みに失敗していることが分かった。一方で経路長が短いときのように、ターゲット記事に近くほぼ最適に経路を選択できる場合は失敗する確率が低いことが分かった。以上より今後のターゲット記事絞り込みとして、中間ゴールがあるという前提の下で、中間ゴールとなる記事の後では最適に動いていると考えられることから絞り込みの方法は本節で述べた方法と同じにし、中間ゴールまでは絞り込みの条件を緩くするか、そもそも行わないという方法で行うことを検証したい。

6. 結論

目的に向けて行動する人の行動履歴から目的推定を行うことの達成を研究目的とし、人の行動原理として「人は最終的な目的の前に中間ゴールを経由することを目指す」と仮説をたてロジスティック回帰モデルを作成し検討を行った。まずモデルを作成するために仮説を示す指標を、人の行動を抽象化していると考えられる Wikispeedia を用いて検討し、Shortest Path Length と Outdegree, PageRank の三要素が仮説を示す指標として利用できることを確認した。そしてこの要素をロジスティック回帰モデルに用いて Wikispeedia を行う探索者の行動履歴から終端記事判別とターゲット記事の絞り込みを行った。結果として終端記事判別で上記の要素が正答率に寄与することが分かった。一方でターゲット記事絞り込みから今後検討する仮説が得られた。

本研究では目的に向けて行動を行う人を Wikispeedia 探索者として抽象化し、人の行動傾向を示す指標を用いたモデルで目的推定を行った結果、目的推定への指標の有効性が確かめられた。この結果は目的が明示的に分かっている人に対しての目的推定になっているが、観察できる目的までの行動履歴は同じになるため、今後の研究として目的が暗黙的である人にも研究で得られた知見を用いることができる。以上では中間ゴールを経由するという仮説の下で検証を行ってきた。今後は目的推定したものからそこに至るまでの行動を提示することについてモデルを作成し検証を行う。また、本研究では中間ゴールの指標として用いた SPL 比較と PageRank 以外の要素について検討する。さらにその後、Wikispeedia 以外のデータセットを用いて人の行動傾向についての仮説を検証する。

参考文献

- Babes, M., Marivate, V., Subramanian, K., & Littman, M. L. (2011). Apprenticeship learning about multiple intentions. In Proceedings of the 28th International Conference on Machine Learning (ICML-11) (pp. 897-904).
- Ho, M. K., Littman, M., MacGlashan, J., Cushman, F., & Austerweil, J. L. (2016). Showing versus doing: Teaching by demonstration. In Advances In Neural Information Processing Systems (pp. 3027-3035).
- Ng, A. Y., & Russell, S. J. (2000). Algorithms for inverse reinforcement learning. In Proceedings of the 17th International Conference on Machine Learning (ICML-00) (pp. 663-670).
- Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: Bringing order to the web. Stanford InfoLab.
- Richard S. Sutton & Andrew G. Barto. (2000) 『強化学習』 三上貞芳, 皆川雅章訳 森北出版.
- West and Leskovec. "Human wayfinding in information networks." In WWW-12. 2012.
- West, Robert, Joelle Pineau, and Doina Precup. "Wikispeedia: An Online Game for Inferring Semantic Distances between Concepts." IJCAI. 2009.
- 五味達朗 (2017). 不整合な演示に基づく半教師あり逆強化学習を介した徒弟学習. 大学院研究年報 理工学研究科篇, (47).
- 牧野貴樹(著, 編集), 澁谷長史(著, 編集), 白川真一(著, 編集), 浅田稔 (著), 麻生英樹 (著), 荒井幸代 (著),..., 吉本潤一郎 (著) (2016)『これからの強化学習』 森北出版.

データセット

SNAP 『Wikispeedia navigation paths』 <https://snap.stanford.edu/data/wikispeedia.html> 2019 年 2 月 21 日アクセス.

連絡先

住所: 〒923-1292 石川県能美市旭台 1-1 北陸先端科学技術大学院大学 日高研究室

名前: 小林瞭

E-mail: kaburobo719@jaist.ac.jp