

Title	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
Author(s)	小林, 瞭
Citation	
Issue Date	2019-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/17564
Rights	
Description	Supervisor: 日高 昇平, 先端科学技術研究科, 修士(知識科学)

修士論文

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

1710082 小林 瞭

主指導教員：日高 昇平

審査委員主査：日高 昇平

審査委員：橋本 敬

DAM HIEU CHI

水元 正晴

北陸先端科学技術大学院大学
先端科学技術研究科 [知識科学]

目次

第1章 序論.....	5
1.1 背景・動機.....	5
1.2 研究目的.....	6
1.3 研究方法.....	7
1.3.1 マルコフ決定過程.....	7
1.3.2 従来研究との比較.....	9
1.3.3 迷路ナビゲーション課題への抽象化.....	10
1.3.4 使用するモデルと分析方法について.....	12
1.4 論文の流れ.....	12
第2章 迷路課題への抽象化と Wikipedia の利用.....	13
2.1 迷路課題と目的推定課題.....	13
2.1.1 迷路課題.....	13
2.1.2 目的推定課題.....	15
2.2 Wikipedia データの利用.....	17
2.2.1 Wikipedia ゲーム.....	18
2.3 本研究のアプローチ.....	19
第3章 Wikipedia データ分析.....	21
3.1 Wikipedia データの構造.....	21
3.2 Wikipedia データ分析準備と仮説.....	23
3.3 行動傾向分析.....	26
3.3.1 最短距離減少傾向の分析.....	26
3.3.2 ページランク減少傾向の分析.....	28
3.3.3 Outdegree 傾向の分析.....	30
3.3.4 Outdegree 傾向と PageRank 傾向の相関分析.....	32
3.4 Wikipedia 被験者のターゲット記事予測.....	33
3.4.1 終端記事判別.....	33
3.4.2 ターゲット記事のピンポイント推定.....	36
3.4.3 ターゲット記事の絞り込み.....	38
3.5 分析結果まとめ.....	41
第4章 総合議論.....	42
4.1 研究目標と研究結果から明らかになったこと.....	42
4.2 今後の課題.....	42
第5章 結論.....	43

参考文献.....44

目次

図 1	MDP 相互作用	8
図 2	マルコフ決定過程を現実に応用した一例	8
図 3	道案内の例での従来システムとの比較	9
図 4	二次元格子状迷路	13
図 5	探索者 A の見えている迷路構造と 1 回目遷移の表示	14
図 6	ターゲットに到達した探索者 A とその行動履歴の表示	14
図 7	一部のマス報酬値を知る探索者 B が見えている迷路構造	15
図 8	探索者 C の見えている迷路構造とナビゲータの見えている迷路構造	15
図 9	ターゲットマス別の迷路構造の一例	16
図 10	Wikispeedia リンクの例	22
図 11	スタート記事からターゲット記事までのステップ数別試行数	22
図 12	一試行の総遷移ステップ数別 SPL 値変化	26
図 13	遷移ステップ長さ別記事絞り込み比較	27
図 14	一試行の総遷移ステップ数別 SPL 記事絞り込み失敗比較	27
図 15	遷移成功試行の各記事 PageRank とランダム遷移記事 PageRank の比較	29
図 16	総記事数別 Outdegree 比較	30
図 17	遷移成功試行の各記事 Outdegree とランダム遷移記事 Outdegree の比較	31
図 18	PageRank と Outdegree の相関比較	32
図 19	終端記事判別に用いる学習データ	33
図 20	終端記事判別正答率（窓に含まれる記事数で統合）	34
図 21	終端記事判別正答率	34
図 22	Outdegree の回帰係数重み比較	35
図 23	PageRank の回帰係数重み比較	35
図 24	ピンポイントなターゲット記事名推定に使用する学習データ	36
図 25	経路長別のピンポイントターゲット記事平均正答率	37
図 26	累計記事推定成功率	38
図 27	記事推定成功回数（5 回遷移）	39
図 28	累計記事推定成功率（5 回遷移）	39

表目次

表 1	現実世界の人と迷路課題の探索者の対応関係.....	11
表 2	現実世界でのナビゲータと迷路課題のナビゲータの対応関係.....	11
表 3	マルコフ決定過程, 迷路課題と Wikipedia の対応関係.....	18
表 4	Wikispeedia データ構造.....	21
表 5	人の行動仮説と Wikispeedia での被験者行動仮説.....	25
表 6	学習データ記事と高順位で推定された記事の比較.....	40

第 1 章 序論

1.1 背景・動機

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する。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るため、実際には存在する選択肢の価値に気づかず、それら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だろう。つまり、自分自身で思いつ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の中に最良の行動があると思い込んでしまうことが多いのである。

このような我々の様子は「旅人」に似ている。旅人はある目的地を目指して、初めて訪れる場所を進んでいく。このとき、旅人は目的地へ通ずる最良の道（ここでの“良さ”には目的地までの近さの他にも、道の歩きやすさなどが複合的に考慮されるだろう）を選ぼうとするが、時には道についての知識不足から最良ではない道を選んでしまうこともある。このような旅人に対してより良い道を知るきっかけとなるのは、道について知識のある「案内人」の情報である。例えば、旅行者の情報よりも地元の人の方がその土地についてより詳しい情報を持っていることを考えると想像しやすいだろう。結果として、案内人から情報を得ることで旅人が最良の道を選択できる可能性を増やすことができるのである。

旅人の場合は以上のように案内人から情報を得ることで目的地へ向かう最良の道を知ることができる可能性がある。それでは、現実の我々にとっての「案内人」がいれば最良の行動選択肢を知る機会が作れると考えた。よって本研究では我々一人ひとりに合わせた「案内人」を作ることで、最良選択肢を知る機会を増やすことを研究の動機としている。ここで本研究での「案内人」とは人に対して「人が選択できる方法のうち目的達成に最も有効な行動の提示を行う存在」とする。さらに「案内人」は人が行動の提示を求めたかどうかに限らず、いつでも行動の提示を行うことを考えたい。これは人によっては余計なお世話となってしまう可能性があるが、もし人が計画していた行動選択肢以外の行動を「案内人」から提示されれば、その人にとって選択できる方法が増えたことになり、増えた選択肢の中には計画していた方法よりも優れた方法が含まれているかもしれないからである。現実の例で考えてみよう。例えば、健康になりたいためヘルシーな食事をとりたいと思う人とその人のアドバイザーがいるとする。あるとき、彼がヘルシーな料理を提供していると思う店 A へ行こうとしたとする。また一方で、ヘルシーとは限らないが健康になることで評判の店 B が新装開店したとしよう。そうするとまず、彼が店 A へ食べに行こうとしたことをアドバイザーが察知し、そこから彼の目的「ヘルシーな食事をとりたい」を予想する。さらにアドバイザー

一は「健康になる食事をとりたいからヘルシーな食事を望んでいる」と目的を先読みして店 B を紹介する。こうすることで、店 B を提案された人は、A 店よりも B 店の方が目的の「健康になりたい」を満たすと気が付けるかもしれない。もし新規開店の情報をアドバイザーから得られなければ、B 店を知るのは先のことになってしまうだろう。

1.2 研究目的

本研究では、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を研究目的とする。このとき、人にとって必ずしも選択肢として明示されていないとしても、最良の行動選択肢であれば提示することを考える。ここで、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何の目的に向けられているかに依存するため、その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を明示的に入力して行動を検索してくれれば良いが、先述の通り、自身の持つ行動選択肢が最良だと思いつく場合など、行動選択者が自覚的に行動の目的を明示することは難しいと予想される。そのため、ここでは人の目的を推定することから行動選択肢の提示まで考えていく。以上より本研究では、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を研究目的とする。

以上の内容を研究目的にするにあたり本研究では人の行動選択について「人は目的達成に最適な行動を選択しようとしている」ことを前提とする。この前提は非常に強い縛りではあるが、問題を簡単にとらえるための一般的な手法であり、人の選択モデルとして行動経済学でも同様に「人は合理的かつ功利的な判断のもとに動くものである」という前提がよく用いられている。もちろん実際には人は目的達成に関係のない行動や非合理的な行動であっても選択してしまうことがあるため、研究目的としているシステムを現実へ適用するときにはこの前提を外れた行動もあり得ることを考慮する必要がある。なお、以上の前提では人が最適な行動を選択しようとしているだけであり、実際に最適な行動を選択できたかどうかは限定しない。これは 1.1 節で述べた通り、人が思いつける行動選択肢はその人の認知能力に限定されたものであり、必ずしも最適な行動選択肢が含まれるとは限らないからである。

1.3 研究方法

本研究ではある人の行動遷移の履歴に基づきその人の目的推定を行うことを目指す。行動遷移の履歴を目的推定に用いる理由は、行動とは目的を見据えて選択されるものであるため、その行動に目的が何らかの形で反映されると考えたからである。具体的に人の行動履歴には、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があると仮定する。このように仮定することで人の行動履歴の時間的変化からそれに反映されている潜在的な目的に関する情報を得ることが可能になる。一方で、人は状況に関する不完全な知識に基づき、必ずしも最適とは限らない行動を選択しながら、目的を果たすための一連の行動の履歴を生成すると仮定する。以上の仮定は、一般にマルコフ決定過程(後述)として定式化できる。その上で、ある種のマルコフ決定過程から確率的に生成された行動履歴の標本からそこに潜在する目的を推定することを目指す。1.3節ではマルコフ決定過程への定式化について述べ、マルコフ決定過程で定式化したことによる従来研究との比較を行う。さらにマルコフ決定過程を前提として具体的な課題として、迷路課題と迷路ナビゲーション課題を利用することについて説明を行う。最後に抽象化した課題を解くために使用するモデルと分析方法について述べて、第2章で行う人の行動を迷路課題へ抽象化することについて繋げる。

1.3.1 マルコフ決定過程

マルコフ決定過程 (Markov Decision Process : MDP) とは、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様子を表すものである。これは図1のようにエージェントとエージェントを取り巻く環境との相互作用で表すことができる。図1で s_t は状態のことであり、エージェントが行動する環境の変化を指す。また a_t は行動のことであり、エージェントが環境に対して働きかける内容を指す。最後に r_{t+1} は報酬のことであり、ある状態から行動し次の状態に移行したことで環境から与えられる価値を指す。また添え字の t は時間を表している。相互作用の順番として、

- (1) 環境からエージェントに現在の状態情報 s_t を引き渡す
- (2) エージェントが環境に行動 a_t を行う
- (3) 行動 a_t で変化した先の状態 s_{t+1} と、状態 s_{t+1} に至ったことで得られる報酬 r_{t+1} を環境からエージェントに引き渡す

の3ステップが1単位時間で行われる。

このようなマルコフ決定過程を用いることで、人の行動がその時々で変わることを各状態で行動選択肢が変化することで表せ、さらに目的に向けて行動が連鎖していることも行動により遷移する状態が繋がっていることで表すことができる。また目的に向けて行動す

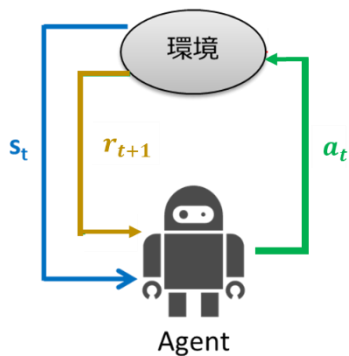


図1 MDP 相互作用

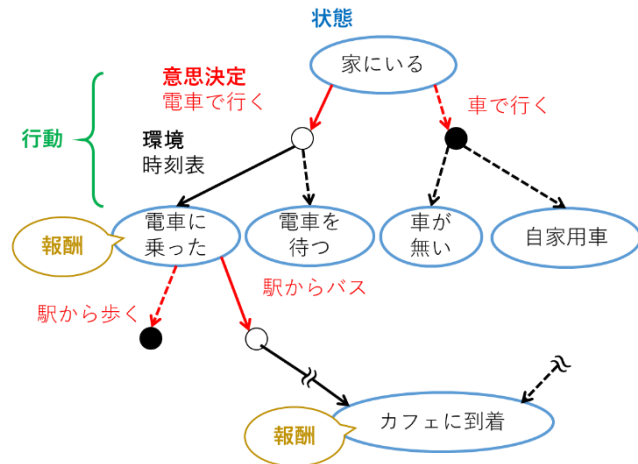


図2 マルコフ決定過程を現実に応用した一例

る様子をメタ的に見ることで、各状態を一つ一つ目的を達成したものとしてもとらえることができる。メタ的に見るかどうかはその時々の問題設定によって変化する。それでは実際にマルコフ決定過程で現実の例の表現が可能であることを簡単な例で示す。図2は現実で、ある人が家にいる状態からカフェに行くまでの行動・状況・行動選択肢の連鎖的依存関係をMDP相互作用として表したものである。青色の丸は状態を表す。緑色の中かっこの範囲は意思決定と環境を合わせた人の行動を表す（意思決定と環境については後述）。そして上から2段目の状態に付いている黄色の吹き出しは報酬を表す。実線の矢印は実際に人が選択したもの、破線の矢印は選択肢として与えられていたが選択されなかったものを表している。図2の例の流れとして、ある人が「家にいる」状態のときに、カフェに行こうと思い立ったとする。このときカフェに行くための選択肢として「電車で行く」「車で行く」の選択肢が現れ、「電車で行く」を選択する。電車に乗れるかどうかは自分ではどうしようもないため決められた時刻表に従うしかないが、例では乗れたとして「電車に乗った」状態になったとする。すると電車に乗ったことで報酬が得られる。このときの報酬の内容は人によって変化し、例えば運転免許を持っていない場合は価値が高くなり、反対に満員電車で嫌な思いをするのなら価値は低くなる。ここまでがMDPでの1単位時間となる。単位時間は環境との相互作用の捉え方により変化するため、より細かくすることも大まかにすることもできる。さらに例を続けて、「電車に乗った」状態からカフェ最寄り駅に着いたとして、「駅から歩く」「駅からバスに乗る」の選択肢が現れ、「駅からバスに乗る」を選択する。そして最終的に「カフェに到着した」状態に至り、初期のカフェに行くという目的からその達成までの行動・状態・行動選択肢の連鎖的依存関係が示せていることが分かる。

1.3.2 従来研究との比較

本研究では目的達成までの方法を提示することを考えているが、これは従来のリコメンドシステムと似ている点がある。従来のリコメンドシステムは協調フィルタリング、内容ベースリコメンド法に大きく分けられる。協調フィルタリングとはアイテム利用者の行動履歴をもとにリコメンドする方法である。例えば Amazon の「こんな商品を買ったひとは、このような商品も買っています」という機能が有名であり、これはあるユーザーと他のユーザーの購買履歴から類似度を計算し、嗜好性の高い商品を提案する。一方で内容ベースリコメンド法は、アイテムの特徴ベクトルからリコメンドする方法である。例えば、レストラン検索で「安い・海鮮」とキーワードを入れた時にその条件に当てはまるレストランを表示する機能である。

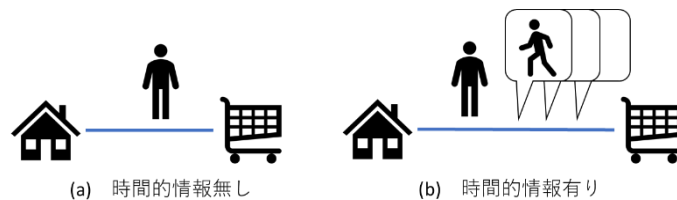


図3 道案内の例での従来システムとの比較

探索者の履歴をもとにその探索者の選好を推定して行動選択肢を提示する点で従来研究と本研究の目指すシステムは同じである。一方本研究で目指すシステムは探索者の行動履歴の時間的前後関係にも注目する。従来のシステムではその人の履歴だけに基づいて行動選択肢を提示する。例えば、スーパーで買い物かごに人参・玉ねぎ・ジャガイモ・鶏肉を入れたときに、これらの材料がカレー又はシチューの材料として一般的だと推定して、おすすめの食材としてカレー粉やシチューの素を提示することができるだろう。一方で従来システムでは十分な提案ができないこともある。例えば道案内をすることを考えて、図3のようにある人がスーパーから家へ帰るときにある時点で家とスーパーの中間地点にいたとする。このとき従来のシステムに入力される情報は、人が家とスーパーの中間地点にいるということだけであるため、その人にとってもう必要ないのにスーパーまでの道に関する情報を提示してしまうかもしれない。一方本研究の目指すシステムでは、その人の時間的変化も考慮するため、スーパーにいたことと家とスーパーの中間地点にいる、という情報が入力として与えられる。そのため、中間地点から家までの道に限定して情報を提供することができる。

以上より本研究では従来研究と比べて、時間的変化について注目することを説明した。次の節で説明する迷路課題ナビゲーションでは、迷路課題においてターゲットマスまでの経路が可換とは限らないことを利用して、行動の順序によって結果が変わる点に時間的変化があるという意味づけを行う。

1.3.3 迷路ナビゲーション課題への抽象化

1.3.1 節で現実世界での人の行動についてマルコフ決定過程を用いて表すことを述べた。しかし現実の課題には様々な要素を考慮する必要があるため分析が難しい。そこでまずは問題について考えやすくするために、マルコフ決定過程の特徴を持ちつつ簡略化された迷路課題について考える。迷路課題とはスタートからゴールまでを繋ぐ経路を作る課題である。本研究では目的に向けて探索的に行動する人に対して、行動選択肢を提案するようなナビゲータを作ることを研究目的としていたため、探索を行う「人」と「ナビゲータ」の関係を迷路課題では「迷路課題探索者」と「迷路課題ナビゲータ」の関係とする。そして本研究が目指すシステムが解く課題について、迷路課題を用いた表現として「迷路課題を解く探索者に対して価値が高くなる行動を提案する課題」とする。また、これを簡略的に「迷路ナビゲーション課題」と呼ぶ。

迷路課題への簡略化に伴って、本研究が注目している現実世界の特徴が迷路課題でも特徴として表れていることを説明する。現実世界と迷路課題の対応関係として次の点が挙げられる(表1)。1つ目は、迷路にはスタートマス、ターゲットマス、どちらでもないマスの3種類が存在すること。これは現実世界でもある時点をもっとも最初の状態とすれば、目的とする状態と目的状態に至るまでの状態として表せている。2つ目は、迷路中の各マスに報酬値が設定されていて、あるマスへ訪れることで報酬が得られ、その報酬値は探索者の目的に応じて変化すること。これは現実世界でも最終的な目的に対して効果的な状態とは何か決まる点で同じ意味を持つ。3つ目は、迷路課題がターゲットマスまでの遷移で得られる報酬値総和で一連の行動の評価を行うこと。これは現実世界では目的状態にたどり着いたときに何をもって達成したとみなすかどうかの評価を行うこととして表している。4つ目は、探索者に与えられている情報がスタートマスとターゲットマスの位置のみで、各マスの報酬値はそこへ遷移しないと分からないこと。これは現実世界でも、実際に行動してある状態にならないとその状態の良さが分からないこととして同じ意味を持つ。5つ目は迷路課題では探索者の行動は、行動を重ねることで次のマスへ進んでいること。現実世界でもある時点毎に選択された行動を重ねている。以上の点で一般的な迷路課題と現実世界の対応関係がみられることを説明できる。

次に迷路課題ナビゲータと現実世界のナビゲータの対応関係について設定したものを説明する(表2)。1つ目は、迷路課題のナビゲータはターゲットマスに向けて遷移している途中の探索者に、報酬を増やしつつターゲットマスへ至るまでのナビゲーションを行うことを目的としている。本研究で目指す現実世界のナビゲータも目的状態に向けて行動する人に、その人にとって良いとされる状態に至る行動を提案する。2つ目は、迷路課題ナビゲータへの問題設定として、ナビゲータに与えられている情報が探索者の遷移履歴と各マスの報酬値のみであり、探索者のターゲットマスと探索者の方策はナビゲータに与えられて

いないこと。現実世界のナビゲータでも、与えられている情報は人の行動履歴と各状態のもつ価値だけであり、人の目的やその人がもつ価値観については与えられていない。

表 1 現実世界の人と迷路課題の探索者の対応関係

現実世界	迷路課題
初期状態	スタートマス
目的とする状態	ターゲットマス
行動の選択肢数	次に遷移可能なマスの数
行動の系列	マスを遷移した履歴
状態の良さは目的による	マスの報酬値は課題設定による
状態に至ることで初めて フィードバックを受ける	マスを訪れることで報酬が得られる
人は目的地到達後に初めて 今までの行動の良し悪しが分かる	ターゲットマスへの遷移で初めて 遷移したマスの報酬値が得られる
良いと判断される状態を 多く遷移するほど一連の行動が 良かったと評価できる	報酬値の高いマスを 多く訪れるほどターゲットマスまでの 行動が良いと評価できる
人は行動を重ねることで状態を 更新していく	探索者はマスの遷移を重ねることで 新たなマスへの遷移を行える

表 2 現実世界でのナビゲータと迷路課題のナビゲータの対応関係

現実世界	迷路課題のナビゲータ
人に行動を提案する	ナビゲータは探索者にあるマスへの 遷移を提案する
人の行動履歴情報を持つ	探索者の遷移したマスの履歴情報を持つ
状態の価値情報を持つ	各マスの報酬値情報を持つ
人の目的は不明	探索者ターゲットマス不明
人のもつ価値観は不明	探索者の価値関数と方策不明

1.3.4 使用するモデルと分析方法について

本研究ではナビゲータが探索者に行動の提案をするために、まず強化学習モデルを使用し、探索者の行動の様子を数理的に表現する。強化学習とは未知の環境から各状態に対応する報酬関数のみを所与として、その累積報酬を最大化する行動を獲得するための手段として Sutton や Watkins によって考案された機械学習の手法である。強化学習の詳細については第2章で述べる。次に、探索者の目的が推定できた場合に、その目的へ向けた最適な行動選択肢を提示するために逆強化学習モデルを使用し、提示が可能であることを示す。逆強化学習とは行動データや状態データから報酬を推定する学習手法で、強化学習で「状態」や「報酬」の設定が難しいことへの解決策として Ng & Russell(2000)によって考案された。ただし、本研究で用いた Wikipedia データ（後述）はデータ数が少なく、一般に大量のデータを必要とする逆強化学習モデルは利用できないため、Wikipedia データに対して逆強化学習モデルの適用は考えない。逆強化学習の詳細についても第2章で述べる。最後に探索者の目的推定を行うために、ロジスティック回帰分析を使用して人の行動を説明する複数の変数について検討した。回帰分析とは目的変数を予想するための説明変数を使って推論の手がかりとする分析方法であり、説明変数が一変数であれば単回帰分析、二変数以上であれば重回帰分析と呼ばれ、ロジスティック回帰分析はロジスティック関数に説明変数と重みの式を代入している。

1.4 論文の流れ

本論文では、まず人の行動系列は意思決定の連鎖であるという仮定と迷路課題ナビゲーションへ抽象化して考えることについて述べた。第2章では、意思決定の連鎖であるという仮定を踏まえて迷路課題と目的推定課題について述べ、強化学習を用いて研究目的としている目的推定が可能であることを述べる。また、具体的な研究題材として用いた Wikipedia データの内容と取り上げる理由について述べる。Wikipedia データは以降に続く第3章の分析や議論の題材になっている。第3章では目的推定法の提案を目指して Wikipedia データの回帰分析を行った。第4章では序論で提示された目的に対して実験で明らかになった点を述べ、今後の課題を示す。最後に第5章で本論文の結論を述べる。

第 2 章 迷路課題への抽象化と Wikispeedia の利用

1 章で人と案内人の関係を迷路課題探索者と迷路課題ナビゲータの関係へ抽象化することについて説明した。本章では人の行動を迷路課題へ抽象化したときの課題設定について説明し、迷路課題という理想化された課題で研究目的である目的推定や行動選択肢の提示が行えることを確認する。2.1 節で人の行動連鎖を迷路課題へ抽象化することと、案内人の行動選択肢提示を行うための最初のステップとして、目的推定課題へ抽象化することについて述べる。2.2 節では 2.1 節で抽象化された課題が、実データとして利用できる Wikispeedia でも同じ意味を持つことを説明し第 3 章へ繋げる。

2.1 迷路課題と目的推定課題

2.1.1 迷路課題

迷路課題とはスタートマスからターゲットマスまで経路を辿る課題である。経路を辿る人は探索者と呼ばれる。迷路中のマスには報酬値が設定されているが、探索者はそのマスへ訪れた時点では報酬を得られず、ターゲット地点に到達したときに初めて各マス報酬値の総和が探索者に与えられる。以下では迷路課題の説明と迷路を解く探索者について説明する。迷路構造はその迷路課題によって決定されるが、ここでは図示を簡単にするため二次元格子状迷路(図 4)として表す。図 4 の迷路構造は 4×4 マスの格子状で、探索者は縦横一マスずつ隣接マスへ遷移できるが、 4×4 の外には出ることができない。左下マスの S はスタートマス、右上マスの T はターゲットマスを指し、この例ではそれぞれ一つずつ示されている。探索者は S から T まで遷移することを求められる。また、報酬値が異なるマスが

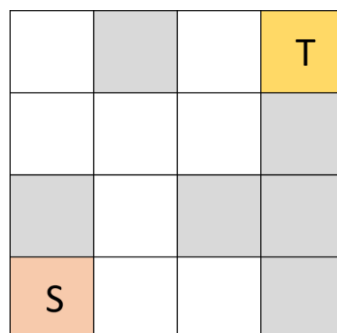


図 4 二次元格子状迷路

2種類存在し、白色マスは灰色マスよりも報酬値が高いマスである。ただし二次元格子状迷路にすることで遷移先のマスが最大4つに限定されていることに注意する。

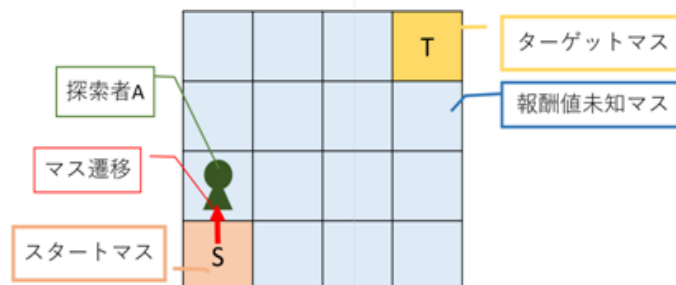


図5 探索者 A の見えている迷路構造と 1 回目遷移の表示

この迷路を探索者が探索する時、例えば上に 1 マス動いたとすると図5のように表せる。図5は探索者 A にとって迷路構造がどのように見えているのかを表している。図4で白色と灰色で表されたマスが図5では水色の報酬値未知のマスとして表示されている。これは探索者 A から見て迷路中の各マスに与えられている報酬値が未知であることを表している。探索者 A が報酬値を知るためには一度そのマスを通らなければいけないが、通ったマスの報酬値はターゲットマスに到着したときにその総和で与えられるため、各マスの報酬値については分からない。このまま探索者 A がターゲットマスまで進んだとすると図6のように表せる。

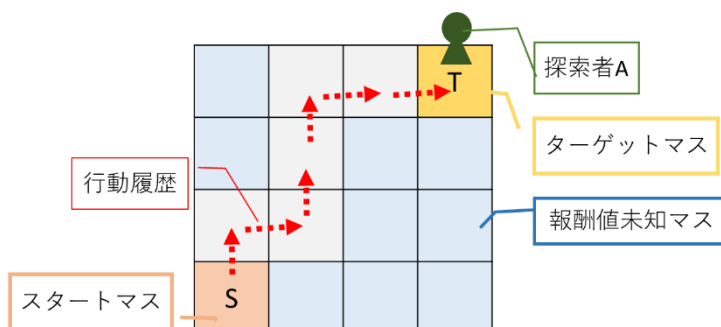


図6 ターゲットに到達した探索者 A とその行動履歴の表示

図6では探索者 A はスタートマスからターゲットマスまで赤色の矢印で示される経路進んだことを表している。ターゲットマスに到着したとき探索者 A は通ったマスの報酬値の総和が得られるが各マスの報酬値については未知である。図4で示された迷路構造の報酬値設定の場合、報酬値が低いマスも高いマスも両方経由しているため、探索者 A から見ると各マスの報酬値は平均的に淡い灰色で表されている。

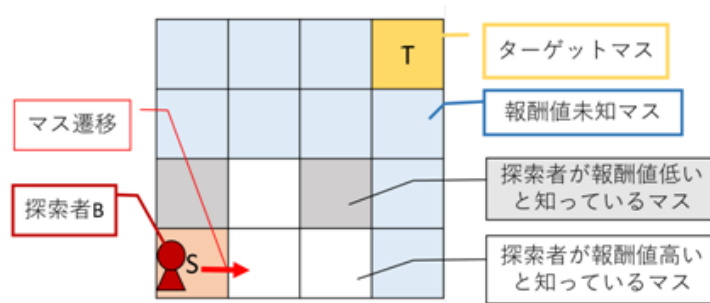


図7 一部のマスの報酬値を知る探索者 B が見ている迷路構造

ここで、現実世界を抽象化していることを考えると、ターゲット記事までの迷路中のマスの報酬値について全く知らないというのは不自然かもしれない。そのようなときは図7のように表せる。図7では探索者 B がスタートマスとターゲットマス以外のマスの報酬値あらかじめ知っていることを表している。このように探索者によって迷路構造に対する知識の違いが生じる場合がある。

2.1.2 目的推定課題

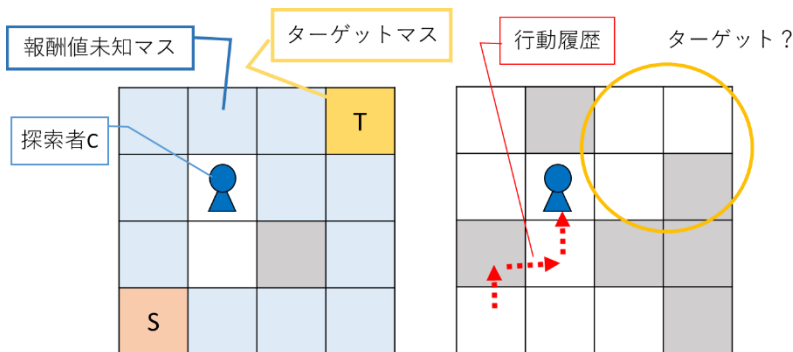


図8 左：探索中の探索者 C と探索者 C の見えている迷路構造

右：ナビゲータに与えられる探索中の探索者 C の行動履歴情報とナビゲータの見えている迷路構造

目的推定課題とは、エージェントに行動のナビゲーションをするためにエージェントの終端状態を推定する課題である。推定にはエージェントの終端状態到達前までの行動履歴とエージェントの環境情報を既知のものとして用いる。目的推定課題は迷路課題に抽象化した場合、探索者のターゲットマスを推定することと同義である。ここで具体的に2.1.1節と同じく二次元格子状迷路に抽象化した迷路構造の一例を用いてターゲット推定について説明する。図8の左図は探索者 C がターゲットマスに向けてマスを遷移している様子と探索者 C が迷路構造内の報酬値マスを一部知っていることを示している。図8の右図はナビゲータが探索者 C の遷移履歴からターゲットマスを大まかに推定している様子と、そのときナビゲータからは迷路構造内各マスの報酬値が分かっていることを表している。

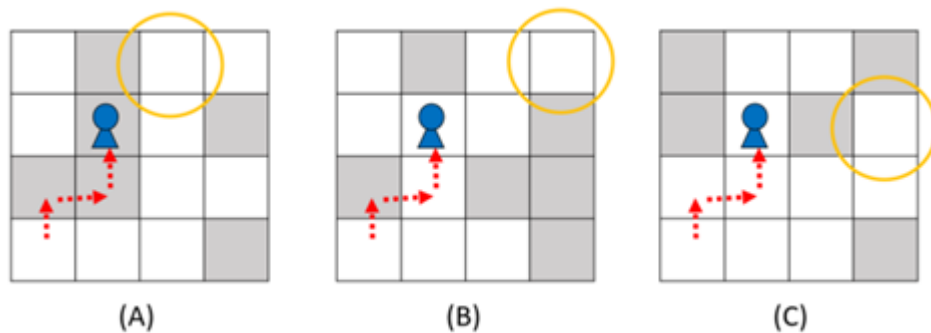


図9 ターゲットマス別の迷路構造の一例

図8の例での目的推定課題は、探索中の探索者Cのターゲットマスをナビゲータが推定することである。推定を行うときにナビゲータに与えられている情報は、(1)探索者Cのマス遷移履歴、(2)探索者Cの現在地点、(3)迷路構造内各マスの報酬値である。目的推定課題の難しい点は(3)迷路構造内各マスの報酬値配置について何種類もある中から推定する必要がある点である。例えば図8の右図ではナビゲータは1種類の報酬値配置について見ているが、実際は図9のようにターゲットマスによって迷路構造内の報酬値配置が異なる。これは第1章の研究目的で触れているように、同じ行動を選択したとしてもその人の目的によって、その行動が目的にとって良いものか悪いものか変化することを表している。目的推定課題は特に、この迷路構造がターゲットによって変化することに注目して取り組む必要がある。

2.2 Wikispeedia データの利用

2.1 節ではマルコフ決定過程を迷路課題や目的推定課題として抽象化したときに、どのような問題設定になるのか説明した。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知識不足に起因する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、実験に工夫が必要になる。そこで、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を反映しているデータセットとして Wikispeedia データ (West, Pineau & Precup, 2009) を利用することを考える。Wikispeedia とは Wikipedia 記事をリンクで遷移する課題であり、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。詳しい課題設定は 2.2.1 で述べるが、Wikispeedia はリンクを辿る事でターゲット記事へ着くことが求められているため、人が想像するリンクと実際に Wikipedia 記事を繋ぐリンクに乖離があるなかでも課題の達成を目指す探索者のことを、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様子にとらえることができる。Wikispeedia データとはこのような課題に対して実際に人が記事間のリンクを辿ったときの記事遷移について記録したものである。Wikispeedia データはもともと、記事リンクを意味ネットワークと考えた時に、その記事を人が遷移するときその人が持つ意味知識が反映されると考えた West らによって作成されたものである (West, Robert, Pineau & Precup, 2009)。以上より、本研究でこの Wikispeedia データを取り上げる理由は、(1) Wikispeedia はマルコフ決定過程の一種であり、かつ(2) 人の一般的な意味知識と Wikipedia のグラフ構造には関連はあるが、それだけでは課題に対して最適な行動をとるのには不十分である、という適度な事前知識を持った探索行動のデータとなっているためである。これは第 1 章でも触れているような、人が思い込んでいる行動選択肢を基に人は行動を選択するが、その選択肢は実際に存在する行動選択肢と異なる場合がありうる状況をモデル化したデータであると解釈できる。人はこの課題に関する不完全な知識を基に目的に向けた行動を選択するため、そうした状況でどのような探索行動をとるか解明できれば目的推定の機序解明へとつながると考えられる。続く第 3 章では Wikispeedia データのページ遷移系列から遷移する人の不完全な知識に基づく意思決定を特徴づける要因を検討するために分析を行う。

2.2.1 Wikispeedia ゲーム

Wikispeedia ゲームではまず被験者に一度の試行の中で、ある 2 つの Wikipedia 上のページが一方はスタート記事として、もう一方はターゲット記事として与えられる。そして、スタート記事から Wikipedia のハイパーリンクだけをたどり（検索やタグを使ってはいけな
いがブラウザの“戻る“ボタンは使用できる）、ターゲット記事まで到着することが被験者の課題である。さらに、スタート記事を 0 ステップ目の遷移記事、スタート記事からハイパーリンクで遷移した先の記事を 1 ステップ目の遷移記事と呼ぶと、被験者にはこのステップ数をできるだけ少ない数でターゲット記事まで到達することが目標として与えられている。

表 3 マルコフ決定過程，迷路課題と Wikipedia の対応関係

マルコフ決定過程	迷路課題	Wikispeedia ゲーム
環境	迷路構造	Wikipedia 記事のグラフ構造
エージェント	探索者	被験者
状態 s	迷路中のマス	Wikipedia 記事ページ
行動 a	マスの遷移	記事中のハイパーリンク選択
報酬 r	各マスに設定された報酬値	各記事に設定された報酬値
終端状態	ターゲットマス	ターゲット記事
初期状態	スタートマス	スタート記事
—	最適探索者の価値が高いマスが探索者にとっては低い	被験者が最短経路の記事に気づいていない
—	迷路中各マスの報酬値に関して不完全ながらも知識をもつ	記事の繋がりについて被験者の常識を利用できる

第 1 章で述べたマルコフ決定過程 (MDP) と 2.1 節で説明した迷路課題，さらに本研究で人の行動データとして扱う Wikispeedia で用いる言葉の対応関係を表 4 に示す。対応関係は大きく分けて以下の 4 つである。まず 1 つ目に MDP ではエージェントを取り巻く世界のことを環境とし，迷路課題では迷路構造とモデル化し，迷路を探索する人が探索を行う迷路そのものを指す。例えば 2.1 節で説明に用いた二次元格子状迷路もその一つである。MDP の環境は Wikispeedia ゲームでは Wikipedia の記事を点として記事を繋ぐハイパーリンクを辺として見たグラフ構造としてとらえる。2 つ目に，MDP の環境，迷路課題の迷路構造，Wikispeedia ゲームのグラフ構造のそれぞれで，目的に向けて進む存在を，エージェント，探索者，被験者と呼びわける。3 つ目に MDP ではある状態からエージェントが行動を行うことで次の状態へ変化し，同時に報酬がエージェントへ与えられる。迷路課題では迷路中のマスを探索者が遷移し，各マスに設定された報酬を受け取る。Wikispeedia ゲームでは

Wikipedia 記事のページからハイパーリンクを辿って次の記事へ遷移し、各記事に設定された報酬値を得る。このとき、MDP と迷路課題、Wikispeedia ゲームに共通して、報酬の設定方法そのものが目的を表現する手段となる。例えば迷路課題の場合、探索者に通ってもらいたいマスがあればそのマスの報酬値を迷路作成者が上げることで探索者の学習を促すことができる。一方 Wikispeedia ゲームの場合はある記事の報酬値は存在するが、それは課題作成者が設定するのではなく、被験者が設定する点で異なる。本研究では Wikispeedia ゲームの分析者として、被験者が設定した報酬値の推定を行うことで、被験者の行動傾向について解明しようとしている。具体的な Wikispeedia の分析については第 3 章にて検討と議論を行う。最後に 4 つ目に、現実では物事のはっきりとした始まりと終わりの設定が難しいのに対して、分析を行うための数理モデルを考えるために MDP、迷路課題、Wikispeedia ゲームで始点と終点を設定している。

本研究で迷路構造と Wikispeedia ゲームを取り上げた理由は、以下の 2 点で研究対象を抽象化しているからである。1 点目は第 1 章の背景で述べた通り、自身が最適な行動選択肢であると思っけていても実際は他に良い選択肢が存在する様子を表すことができる点である。例えば迷路課題では探索者が学習途中の場合、最適行動を知るエキスパートとは異なるマスへの遷移を行うことがある。これはエキスパート探索者にとって価値が高いマスであっても、学習途中探索者にとっては価値が低く見え、移動すべきマスに気づいていないことを表している。同様に Wikispeedia ゲームでも、被験者がグラフ構造としてターゲット記事までの準最短リンクを遷移することから、最短経路について気づいていないことを表している。2 点目は被験者が Wikipedia 記事のグラフ構造について一部の知識を持っている点である。例えばターゲット記事が「Rainbow」である場合、Wikispeedia データから、多くの被験者が「Color」や「Rain」の記事を遷移していることが確認できる。これは多くの被験者が「Rainbow」と「Color」や「Rain」が結びついていることを常識として知っていて、その常識を Wikispeedia ゲームにも当てはめていることを表している。このような被験者の持つ常識は、迷路課題では探索者が不完全ながらもマスの報酬値を推定して、自身の価値基準とそれに伴う行動基準を持っている様子として表せる。

2.3 本研究のアプローチ

本研究では人の行動履歴からその人の目的を推定し、さらに推定した目的に向けて取るべき行動を提示することを考えていた。第 1 章ではこの問題を考えるために人が目的に向けて行動を選択する様子をマルコフ決定過程の数理モデルに落とし込み、更に簡単に考えるため迷路課題に抽象化した。そして第 2 章では抽象化した迷路課題で探索者の遷移履歴からターゲットマスを推定することについて述べた。次に実際に人の探索行動を参考に目的推定を考えるために Wikispeedia データを利用することを考えた。現実で人が行動を選択

する際にその人が経験してきた情報を統合した「常識」を参考にしているように、Wikipedia被験者も同じく各記事間のハイパーリンクを被験者の常識から大まかに予想してリンクを選んでいる点がWikipediaデータの特徴である。このようなWikipediaデータから人のもつ常識と、行動を決定する価値判断について分析を行うことで研究目的である探索者の目的推定を行うことを考える。具体的にWikipediaデータから(1)記事や記事間の遷移の中で何を価値としているのか、(2)被験者は常識をどのように利用しているのか、(3)どのような判断基準で記事を遷移しているのか、の3点を強化学習での価値関数と方策の代わりとして用いるために分析を行う。Wikipedia分析から被験者が想定している記事の価値と行動の判断基準を知ることができる要素について解明できれば、その結果から被験者の行動モデルを作成し、目的の推定のための情報とすることができる。本研究では人の行動を決定づける候補をWikipediaデータから抽出し、それらを回帰分析することで各候補の重み付けを行う。これは第2章で述べたような、目的ごとの迷路構造をナビゲータが求めようとしていることと同じ意味を持つ。

第 3 章 Wikispeedia データ分析

3.1 Wikispeedia データの構造

本研究では目的推定法について検討するために第 2 章で紹介した 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。データ内容は表 5 にまとめた。

表 4 Wikispeedia データ構造

総記事数	4,604
記事間リンク総数	119,882
記事のカテゴリ構造データ数	5,204
ゲーム成功試行数	51,308
ゲーム失敗試行数	24,875
2つの記事間最短経路長	$4,604 \times 4,604$
強連結記事数	4,051
強連結記事間のゲーム成功試行数	50,834
強連結記事 10 ステップまでの スタート記事の種類数	3,415
強連結記事 10 ステップまでの ターゲット記事の種類数	2,762
記事中の平均ハイパーリンク数	27.6

総記事数とは Wikispeedia で使用した記事の総数であり 4,604 記事存在する。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 10 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 10 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。例えば記事「oil painting」や「photography」は Art という大カテゴリに属している。また記事によっては二つ以上のカテゴリに属していることもある。試行数とは、ある被験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

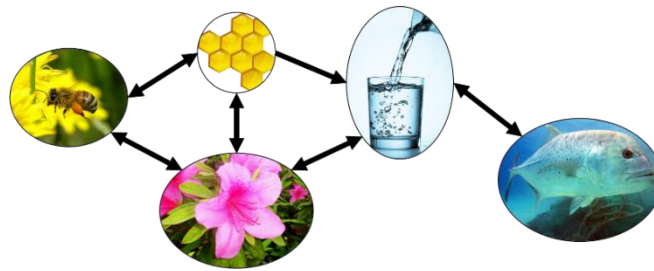


図 10 Wikipedia リンクの例

左上から「Bee」「Honey」「Flower」「Water」「Fish」の記事を指す

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ける。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、Wikispecies 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。

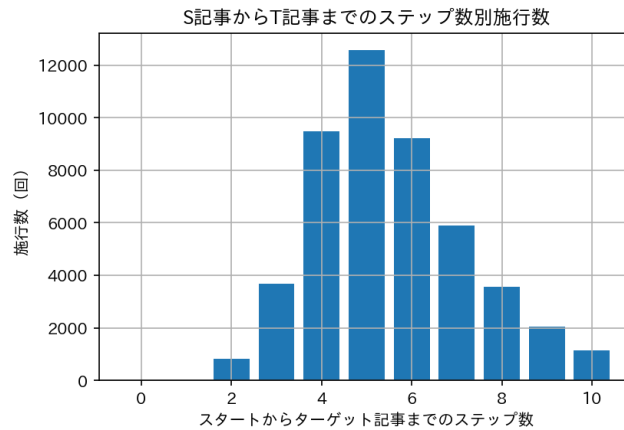


図 11 スタート記事からターゲット記事までのステップ数別試行数

被験者の遷移データでは最小遷移ステップ数 2，最大遷移ステップ数 109 であり，全試行のうち 95% の試行データは記事遷移長さ 10 ステップまで，90% の試行データは 9 ステップまでに含まれる。図 11 は 10 ステップまでにターゲット記事までたどり着いた試行数を示している。10 ステップ遷移内にターゲット記事までたどり着く試行のうち，スタート記事は 3415 記事，ターゲット記事は 2762 記事ある。

3.2 Wikispeedia データ分析準備と仮説

目的推定を行うために Wikispeedia の被験者がどのような記事遷移傾向をもつのか検証する。特に「Wikispeedia 被験者はハブのようなページから狭めていく」ことを仮説として分析を行う。ここでの「ハブ」とは記事がグラフ構造を成していることに関連させたイメージであり、具体的には遷移記事のリンク数や複数の試行で共通して選ばれている度合いなどが「ハブ」を表現する要素であると考えている。以降の分析ではこの「ハブ」を示す要素の検証も行う。例えば「Ant (蟻)」から「Avalanche (雪崩)」まで遷移するときどうするだろうか。一見何の繋がりを持たない2つの記事を繋げるために、できるだけ後々の遷移に繋がりやすい言葉を選ぶだろう。実際に Ant の記事をよく読むと「Earth」や「South_America」などの記事へのハイパーリンクがあり、それらの記事はその他の様々な記事へ繋がっているだろうと予想できるだろう。そして例えば「Earth」に遷移したとすると多くのハイパーリンクの中から「Weather」が選択でき、そこから「Snow」, 「Avalanche」へと無事遷移を終えることができる。

3.1 節で述べた Wikispeedia データからは様々な要素が読み取れる。例えば West らが行った先行研究では Term Frequency – Inverse Document Frequency (TF-IDF), Degree, Shortest Path Length の3つを主要な要素として挙げている。さらに本研究では以上の3つに加え PageRank も要素として加える。それぞれの要素の説明と、それらを現実の人の行動に置き換えたときの意味について以下で述べる。

- TF-IDF とは文書内に出現する単語の重要度を表す。先行研究では Wikipedia 記事の文章について TF-IDF 分析を行っていた。しか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 分析は行わないこととする。
- 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。現実における人の行動に置き換えると、ある状態のもつ選択肢の多さを Degree として考えることができるだろう。
- 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。Wikispeedia ゲームで被験者に求められたことが少ない遷移数でターゲット記事まで向かうことであったことを踏まえて、現実の人の行動に置き換えると、SPL はある状態からその他の状態への最適解となる行動選択肢の判断基準として考えることができるだろう。
- 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。現実の人の行動に置き換えると、多くの人を経由しやすい状態と考えることができるだろう。

なお, Degree, SPL については Wikispeedia データとして与えられているが PageRank については与えられていないため以下のように計算を行った (Page, Brin, Motwani & Winograd, 1999). 計算に用いたプログラムの流れは,

```

 $R_0 \leftarrow S$ 
loop :
 $R_{i+1} \leftarrow AR_i$ 
 $d \leftarrow \|R_i\|_1 - \|R_{i+1}\|_1$ 
 $R_{i+1} \leftarrow R_{i+1} + dE$ 
 $\delta \leftarrow \|R_{i+1} - R_i\|_1$ 
while  $\delta > \epsilon$ 

```

である. ここで S は初期値で Wikispeedia データの記事数 4051 の長さを持つベクトルであり, ベクトル内の総和が 1 になるようにしているためベクトルの中身は全て $1/4051$ である. R_0 はループ 0 回目の PageRank ベクトル (記事数分の長さ 4051) である. A はページから他のページまでのリンク有無を示す行列 (4051×4051) であり, SPL データで最小リンク数が 1 (つまり記事間の直接リンクが存在する) 以外の要素を 0 にした行列である. E はベクトル内容の更新時にベクトルの総和を更新前後で同数にするためのベクトルであり, ここでは $E = S$ としている. δ は PageRank ベクトルの更新値の合計で, この値があらかじめ指定していた閾値 ϵ (今回の計算では 0.001 に設定) 以下になればループを抜け出す.

以上の Wikispeedia 要素を用いて, 続く 3.3 節で検証するために現実の人の行動に対して仮説をたて, さらにその仮説を Wikispeedia 被験者行動に落とし込む. 表 6 は現実の人の行動に対する仮説と, それに対応した Wikispeedia 被験者に対する仮説を一覧で表している. まず 1 つ目に, 人は目的達成に最適な行動を目指すものの知識不足などにより準最適な行動を選択してしまっていることを示すために「人は目的まで最適な行動を選択している」と仮説を立てる. これを棄却することで人が最適な行動を常に選択できているわけではないことを示す. これは Wikispeedia ではターゲット記事までの最短リンク数を減少させることが最適な遷移といえるため, 仮説を Wikispeedia へ落とし込むと「被験者はターゲット記事までに最短経路数が必ず少なくなるリンクを選択する」と表現できる. 次に 2 つ目に, 人は目的に向けた大まかな行動方針として, まずは目的までの確信度の高い行動を選択できるような状態へ向かうと考えていることを示すために「人は最初に大まかに目的に関係ありそうな状態へ向かう」と仮説を立てる. Wikispeedia では大まかに関係のある記事というのは, ある記事からその他の様々な記事への繋がりやすさが高い記事を指すと考えられる. Wikispeedia ではこの繋がりやすさを PageRank として表せる. よって仮説を Wikispeedia へ落とし込むと「被験者は最初の記事遷移では PageRank の高い記事へ遷移する」と表現できる. 3 つ目に, 「人は現在の状態とそれが目的とする状態を繋げる最適行動を計画できないとき, 人は今の状態よりも多くの行動が選択できそうな状態を目指す」と仮説を立てる.

なお、ここでは遷移した先の状態に本当に最適な行動選択肢があるとは限らないものとする。これは Wikispeedia ではスタート記事にターゲット記事までの直接的なリンクが無い場合を、1度の遷移でターゲットに到着できない場合であると考え、さらに今の状態よりも多くの行動が選択できそうな状態を目指すことを、被験者のいる記事よりも多くのリンクを持つ記事へ向かうこととして考えている。よって仮説を Wikispeedia へ落とし込むと「被験者は最初に Outdegree の高い記事へ向かう」と表現できる。

表 5 人の行動仮説と Wikispeedia での被験者行動仮説

	人の行動傾向仮説	Wikispeedia での被験者行動予測
①	人は目的まで最適な行動を選択している	被験者はターゲット記事までに最短経路数が必ず少なくなるリンクを選択する
②	人は最初に大まかに目的に関係ありそうな状態へ向かう	被験者は最初の記事遷移では PageRank の高い記事へ遷移する
③	現在の状態と目的とする状態を繋げる行動を計画できないとき、人は今の状態よりも多くの行動が選択できそうな状態を目指す	被験者は試行の最初に Outdegree の高い記事へ向かう

3.3 行動傾向分析

人の行動を表していると思われる指標についてその傾向を調べる。

3.3.1 最短距離減少傾向の分析

仮説)

人は目的までに取るべき行動系列を全て知っていてその行動を選択している。Wikipedia に置き換えると、被験者が行ったターゲット記事までの経路では、遷移した各記事とターゲット記事の SPL が遷移するごとに下がる。

結果)

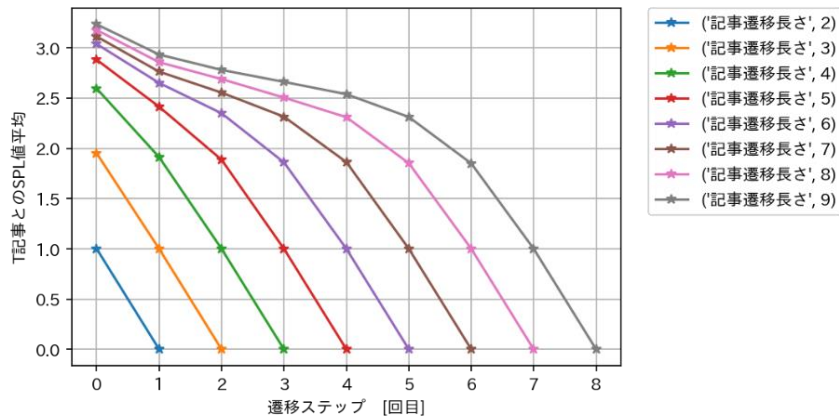


図 12 一試行の総遷移ステップ数別 SPL 値変化

図 12 は遷移ステップ数が 2~9 回でターゲット記事にたどり着いた試行について、各遷移先記事とターゲット記事の SPL を試行回数で平均をとっている。図 12 より、総遷移ステップ数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 の記事をなだらかに減少するよう遷移していることから、必ずしもターゲット記事まで一直線で進んでいないことが分かる。よって上記の仮説は棄却され、人は必ずしもターゲット記事までの最短リンクを辿っているわけでは無いことが分かった。これは人の行動傾向として考えると、スタート記事からターゲット記事の 2 ステップ前までの間では Wikipedia の記事リンク構造と人の思い描く記事リンク構造に乖離があるために SPL の減少がなだらかになってしまっていると考えられる。一方で図 10 の特徴として、どの経路長でも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。このことからターゲット記事へ向かっていること説明にターゲット記事直前の SPL 比較が参考にできると考えられる。

次に図 12 より、人は SPL が少なくとも下がっているか同じである遷移をしていると仮説を立てた。仮説が正しいとすると、被験者の目指すターゲット記事を知らずとも、SPL が大

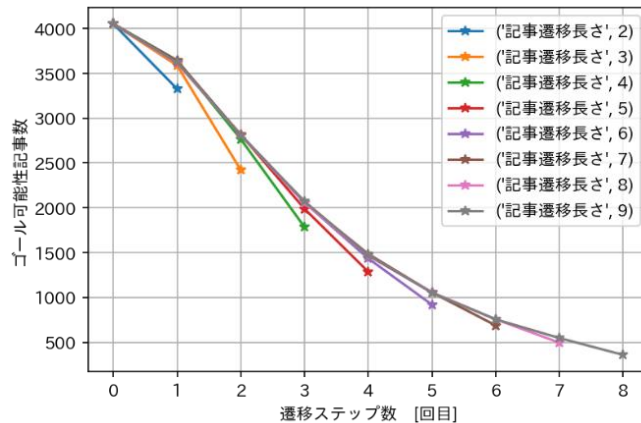


図 13 遷移ステップ長さ別記事絞り込み比較

きくなっている記事をターゲット記事である可能性から除外することでターゲット記事を絞り込むことができるだろう。図 13 は SPL が上がった記事をターゲット記事可能性から除外して記事の絞り込みを行った結果を示している。

図 13 の分析の結果、各経路長で絞り込みができていない確率が異なり、2回、3回経路長では全ての試行（ただし記事戻り無し）で絞り込みが成功していた。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る。以上より、上記の仮説は総記事ステップ数が長くなるほど当てはまらないことが分かった。

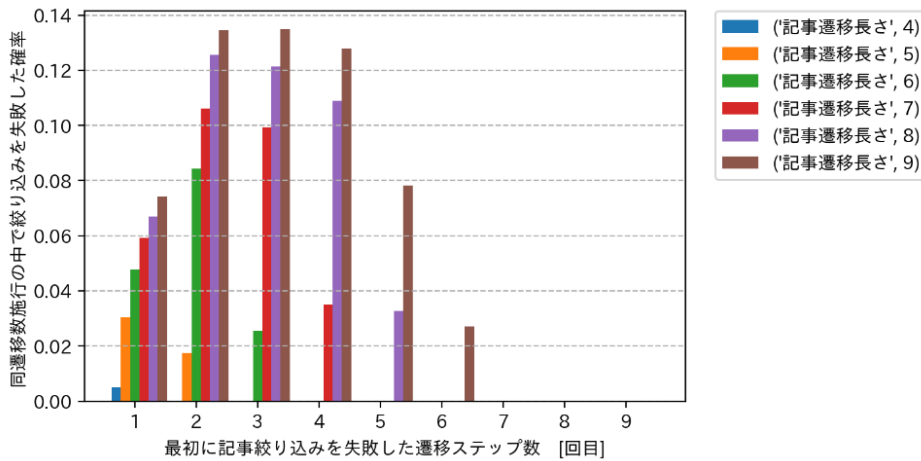


図 14 一試行の総遷移ステップ数別 SPL 記事絞り込み失敗比較

ここで SPL が高くなる記事遷移をしてしまう確率について、記事遷移の回数との関連を分析した。図 14 は記事遷移長さ別に SPL が下がるか同じになり続けるという仮説で、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 14 の結果から、まず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。次に、実験で用いられたスタート記事とターゲット記事は

5 回以内に到達できる記事組み合わせがほとんどを占めているのに対して、経路長の長い試行になるほど 2 回目の遷移で絞り込みの失敗が多いことが分かる。これより、被験者はスタート記事近くの記事遷移では SPL が大きくなってしまふ記事へ向かう傾向もあることが示唆される。たしかに SPL は最小遷移回数を表しているが、人が頭の中で思い描いた記事の繋がりとは異なる繋がりを結ぶことで最小としている可能性があるため、SPL 絞り込みの失敗は、Wikispeedia データのグラフ構造と人の思い描く記事間グラフ構造の差によるものだと考えられる。これより SPL と探索行動の経過に相関があることが示唆されたため、続く 3.4 節での回帰分析に SPL を説明変数として用いる。

3.3.2 ページランク減少傾向の分析

仮説)

人は目的までの大まかな方針を立てて行動すると思われるが、もし自分の現在の状態と目的とする状態までを繋げる直接的な行動を計画できない場合は、まず経路を想定できるようにするために、目的状態まで大まかにでも関係ありそうな状態へ向かう。人にはその行動によって目的状態に向かえることが行動前には分からないが、とりあえず行動して状態を変化させることで目的状態まで行くための何らかの手掛かりを得ようと人は考える。Wikispeedia に置き換えると、PageRank 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のようなものであるため、PageRank の高い記事へとりあえず向かうことで、被験者はターゲット記事まで繋がっていきそうなリンクのある記事を見つけている。

分析内容)

ゲーム成功試行の各遷移記事の PageRank をランダム遷移時と比較する。データはゲーム成功試行である 51,308 試行分の遷移データを用いた。ランダム記事は選択できる記事からランダムに遷移した場合の遷移先 PageRank の期待値である。それぞれのランダム遷移はランダム遷移の直前まで人と同じ遷移をしていた場合を考えている。例えば 2 ステップ遷移目のランダム遷移 PageRank の値は、1 ステップ遷移目までは人と同じ遷移を行い、その遷移した記事から先をランダムにした場合で計算している。

結果)

図 15 はスタート記事からターゲット記事までの記事数が 2 つから 9 つまでの試行について各記事の PageRank を平均したものである。横軸はスタート記事を 0 ステップ目遷移としたときの遷移ステップ、縦軸は遷移した記事の PageRank 平均である。結果としてランダム記事遷移と比較して人の記事遷移には PageRank の大小で傾向がみられることが分かった。図よりすべてのステップ長さで、人は一度目の遷移で PageRank が平均よりも高い記事へ向

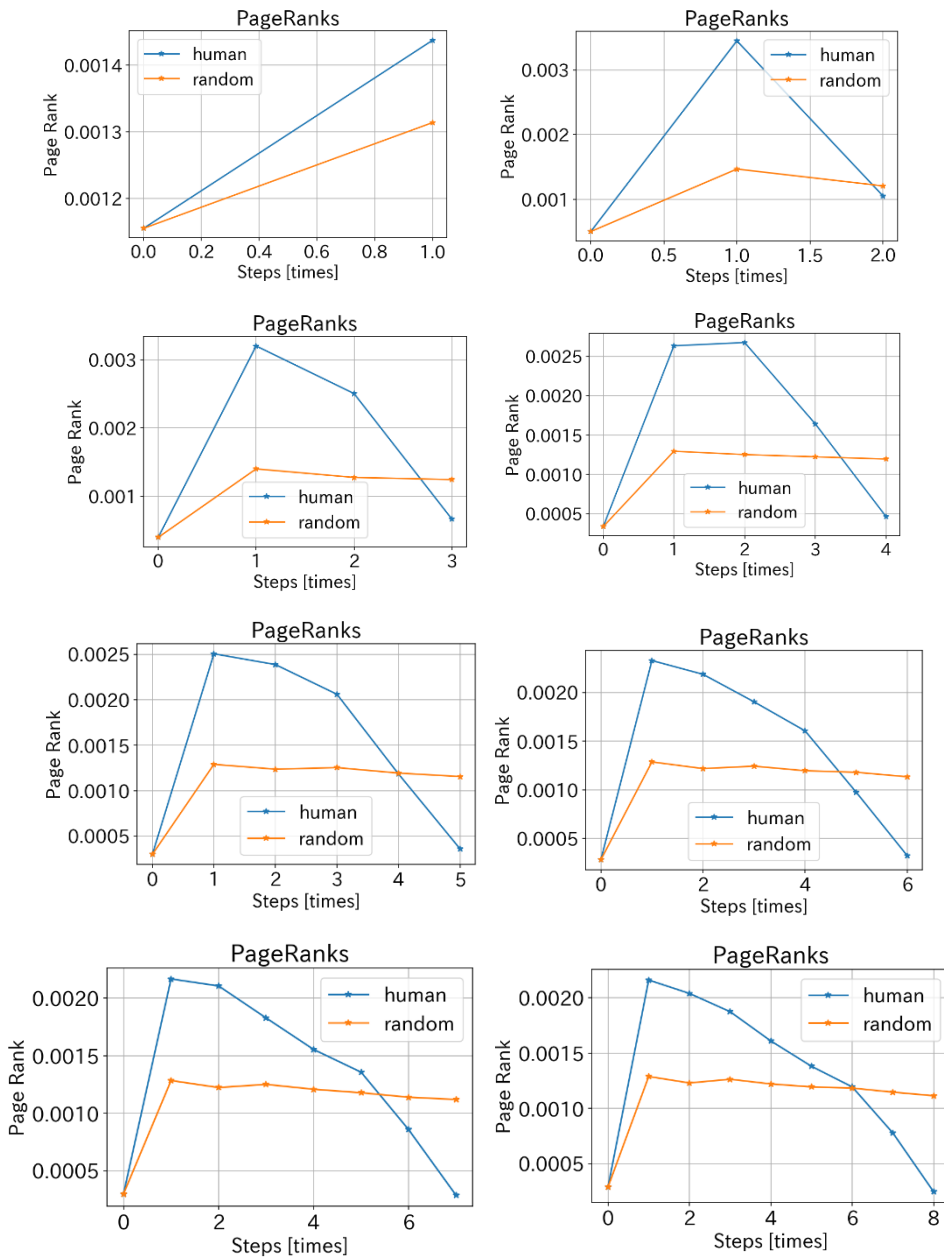


図 15 遷移成功試行の各記事 PageRank とランダム遷移記事 PageRank の比較

かっていることが分かる。また、PageRank が最も高い記事から後の遷移ではターゲット記事まで PageRank の低い記事へ遷移していることが分かる。以上より遷移記事と PageRank の関係から、仮説で述べたように、人は最初に大まかに目的まで関係ありそうな状態に向かう傾向があることが分かった。これより PageRank と探索行動の経過に相関があることが示唆されたため、続く 3.4 節での回帰分析に PageRank を説明変数として用いる。

3.3.3 Outdegree 傾向の分析

仮説)

人は現在の状態とその人が目的とする状態を繋げる直接的な行動を計画できない場合に、その人が取るべき行動を計画できるようにするために、今その人がいる状態よりも選択できる行動が多くなる状態へ向かう。また、言い換えればターゲット記事までたどり着けないリンクしかない記事への遷移を避けている。Wikipedia に置き換えて考えると、Outdegree がある記事から他の記事へのリンク記事数を表していることから、もし被験者がターゲット記事までに最短で繋がるリンクを知らない場合、Outdegree の高い記事へ向かう。なお、本仮説は 4.3.2 節で述べた仮説と似ているが、PageRank では記事遷移を数ステップにわたって行った時の訪れやすさを表していたのに対し、Outdegree では 1 ステップ先の記事遷移に注目している。

結果)

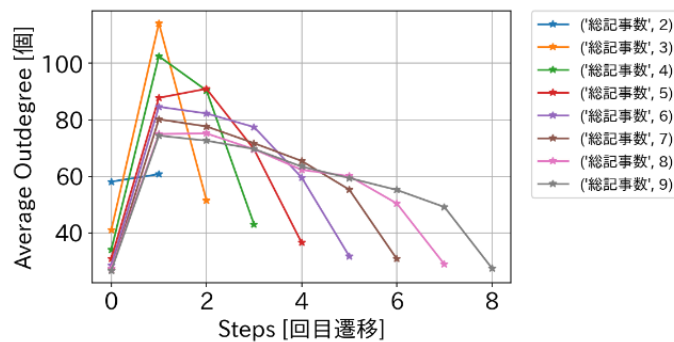


図 16 総記事数別 Outdegree 比較

図 16 は一試行の中でスタート記事からターゲット記事までの遷移で経由した記事数別に、その遷移した記事の順番ごとの Outdegree を平均したものが縦軸で表わされている。横軸はスタート記事を 0 ステップ目遷移記事としてターゲット記事までに遷移したステップ数である。またスタート記事からターゲット記事までの記事数は 2~9 までをそれぞれ表示している。図 16 から Outdegree の 1 回目遷移で Outdegree が多い記事へ遷移していることが分かる。そしてその後はターゲット記事に向けて下がる記事を遷移する傾向がみられる。Outdegree が高い記事へ 1 回目に遷移していることを示すために、チャンスレベルと比較を行う。図 17 はスタート記事からターゲット記事までの記事遷移が成功した試行について、遷移中の各記事の Outdegree とランダムに進んだ場合の記事の Outdegree 平均を、一試行当たりの総記事数 2~9 で場合分けして表している。横軸は、試行ごとのスタート記事を 0 ステップ目遷移記事としてターゲット記事までに遷移したステップ数を表している。

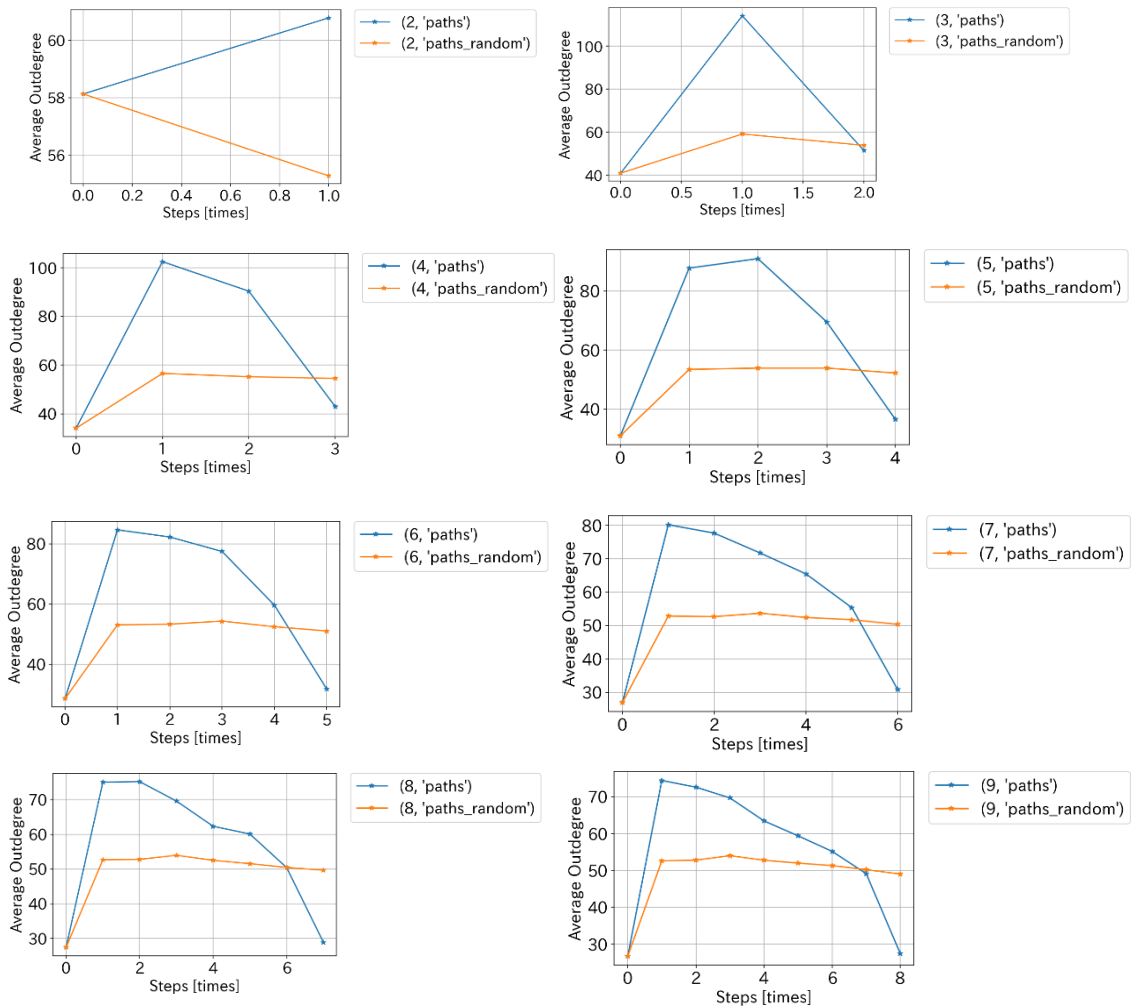


図 17 遷移成功試行の各記事 Outdegree とランダム遷移記事 Outdegree の比較

図 17 から 1 ステップ目遷移でチャンスレベルと比較して Outdegree がより高い記事へ遷移していることが分かる。チャンスレベルよりも Outdegree の高い記事への遷移とは、被験者が他記事へのリンク数多くなる記事へ遷移していることを表していると考えられる。一方でターゲット記事直前から 2 ステップ前の遷移ではチャンスレベルを下回る傾向がみられる。以上より Outdegree の分析からチャンスレベルと比較して最初に値が高くなり、その後はその数を減らす記事への遷移傾向があることが分かった。このことから仮説の通り、人は現在の状態とその人が目的とする状態を繋げる直接的な行動を計画できない場合に、その人が取るべき行動を計画できるようにするために、今その人がいる状態よりも選択できる行動が多くなる状態へ向かうことが分かった。

3.3.4 Outdegree 傾向と PageRank 傾向の相関分析

以上の分析の中で Outdegree の傾向と PageRank の傾向が似ているように見られた。Outdegree と PageRank を検討するとき立てた仮説は異なっているが、被験者の行動傾向の説明に同じ意味をもつのであれば統合したい。そこでこの二つの要素を統合することについて検討する。検討の方法として、横軸を全てに経路長での PageRank、縦軸全ての経路長での Outdegree としてグラフを作成し、これらに相関があることを確認する。相関していることが見られれば二つの指標を統合し回帰分析では PageRank にも Outdegree の傾向が含まれると考える。

結果)

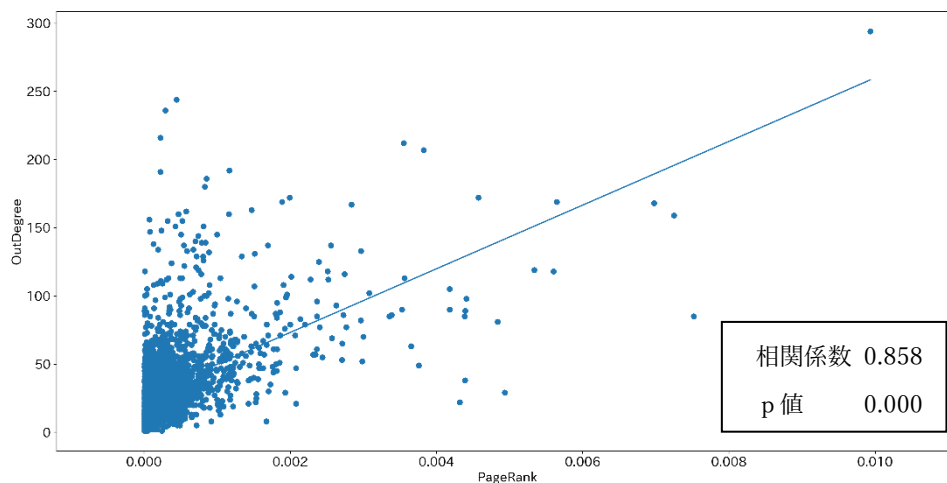


図 18 PageRank と Outdegree の相関比較

図 18 は全てのゲーム試行に対して、試行で遷移した記事ごとの PageRank と Outdegree を表示したものである。直線は各点に対して 1 次近似直線を表示したものである。近似直線の傾きは 23350、切片は 26 で、分散は 911 であった。点の持つ意味として、例えば PageRank が低く Outdegree が高い点は、Outdegree が多いがその記事へリンクしている記事数 Indegree が少ないため PageRank が低くなっていると分かる。PageRank と Outdegree の相関係数は 0.858、p 値は 0.000 であった。PageRank は数ステップの間で記事遷移を考えたときの要素、Outdegree は 1 ステップの記事遷移を考えたときの要素を示すように、それぞれの要素が意味するものは違うが、相関係数の値から PageRank と Outdegree に高い相関があることが分かった。

3.4 Wikispeedia 被験者のターゲット記事予測

Wikispeedia 被験者のターゲット記事予測を回帰分析で行い、正答率と同時にどの説明変数とその正答率に寄与しているのか検討する。説明変数として 3.3 節で被験者の行動経過に相関があることが示唆された SPL と PageRank を用いる。なお Degree と PageRank は異なる仮説のもとで挙げていたが、3.3.4 節の分析から同じ傾向を示していると分かったため、回帰分析では PageRank の傾向に Degree の傾向も含まれるものとして扱う。

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、被験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被験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、被験者のターゲット記事の推定を行う。ターゲット記事推定はまずピンポイントに記事を 1 つ予測した場合の正答率を検討する。次に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。

3.4.1 終端記事判別

使用データ)

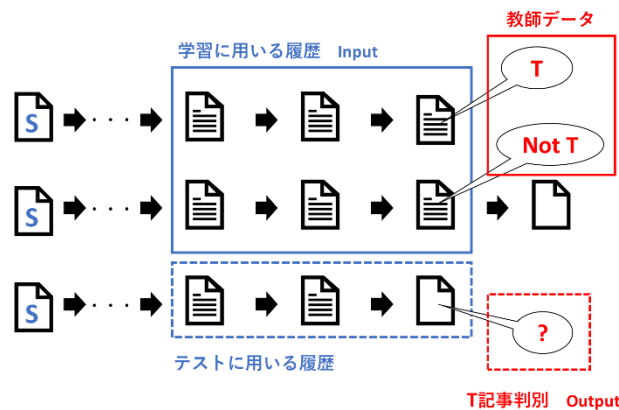


図 19 終端記事判別に用いる学習データ

図 19 は終端記事判別に用いるデータについて説明する図である。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被験者が遷移する。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 19 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 19 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 19 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に選択する。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。

学習を行うシステムに与えられる学習データの Input は、図 19 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 19 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。

結果)

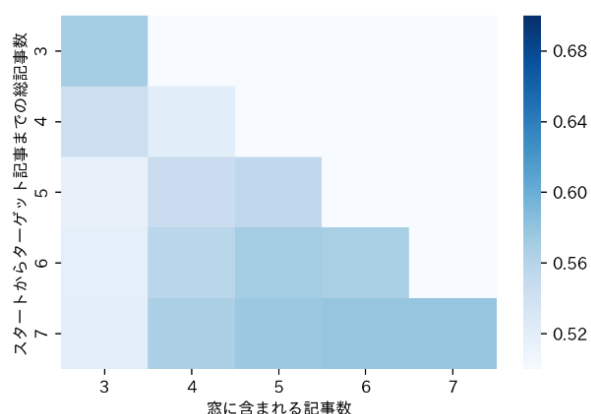


図 21 終端記事判別正答率

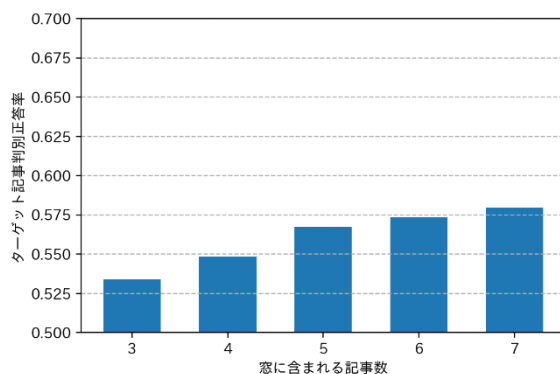


図 20 終端記事判別正答率（窓に含まれる記事数で統合）

図 20 は終端記事がターゲット記事を判断したときの正答率である。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。窓に含む記事数は総記事数以下となるため正答率は出ていない。図 21 は窓に含まれる記事数別のターゲット記事判別正答率を表している。図 20 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。それでは次に窓の大きさ別に見る。図 21 からは窓の大きさと正答率に相関があることが見られる。横軸が 3 のときに低い値をとっているが図 20 で示されているように総記事数が 4～7 であるときの正答率の悪さが原因となっている。図 21 での正答率は全体的にチャンスレベルを超えるものの、高い正答率とは言えない。

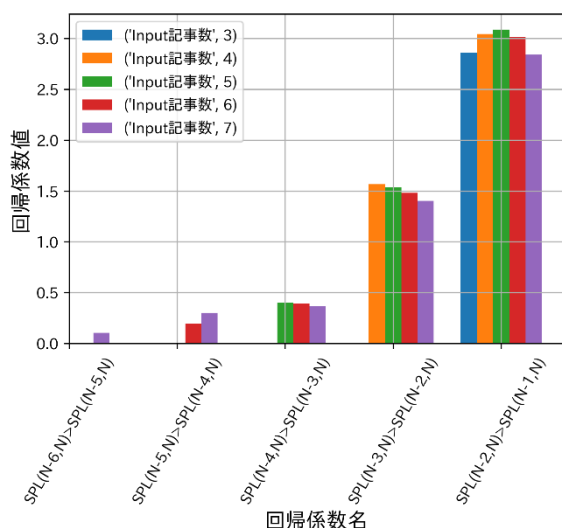


図 22 Outdegree の回帰係数重み比較

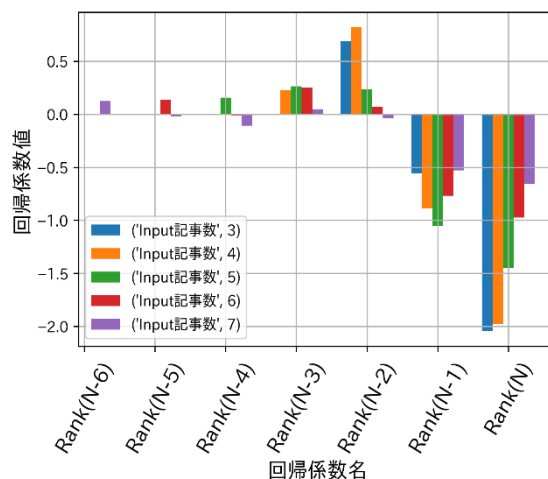


図 23 PageRank の回帰係数重み比較

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 22 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 23 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。図中の N とは終端記事の序数であり、N-1 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 22 から終端記事と 1 つ、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 12 で示しているように、ターゲット記事直前で SPL が急に減少することを示していることが分かる。また、さらに 2 つ、3 つ前記事との SPL 比較も重みが大きいことから、図 12 を見ると SPL 急な減少の開始点であることにもターゲット記事判別に意味があることが分かる。図 23 からは、終端記事の PageRank が最も大きくなるが、Input とし

ての窓が大きくなるほど終端記事の PageRank の重みが小さくなっていることが分かる。これは図 15 の PageRank 分析から、一試行当たりの総記事数が大きくなるとターゲット記事直前の PageRank が下がっている傾向があることが、ターゲット記事に近いことを示す意味を担っていると考えられる。

3.4.2 ターゲット記事のピンポイント推定

研究目的であった目的の推定として Wikipedia 被験者の目指すターゲット記事を 1 記事推定して当てること検討する。

使用したデータ)

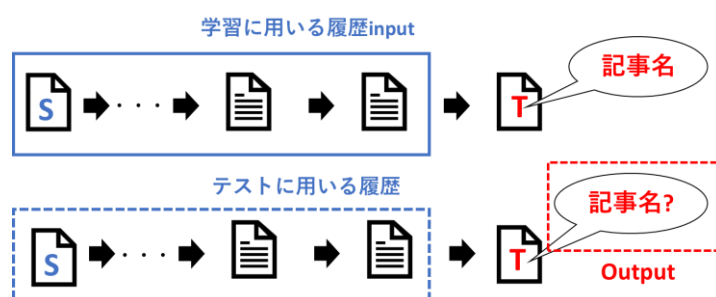


図 24 ピンポイントなターゲット記事名推定に使用する学習データ

図 24 はピンポイントなターゲット記事名推定に用いるデータについて説明する図である。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被験者が遷移する。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である。学習データとして用いる記事 (図 24 実線青枠) を「窓」とするとき、最後に遷移を行った先の記事を終端記事と呼ぶ。学習に用いるデータはその試行のスタート記事からターゲット記事の 1 つ前までの記事全てである。つまり被験者があと 1 ステップでターゲット記事に遷移する場合について分析を行う。学習データさせるときに使用するデータは、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別に作成する。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 の場合、窓に含まれる記事数はターゲット記事を除いた 5 つとなる。

学習を行うシステムに与えられる学習データの Input は、図 24 の実線青枠の窓に含まれる記事に対して、SPL と PageRank 値を計算したものであり、教師データとしてその試行のターゲット記事名が与えられる。SPL 比較は $SPL(u_i, u_{last}) \geq SPL(u_{i+1}, u_{last})$ から True で 1, False で 0 を返す。このとき u_i とはスタート記事を 0 番目として整数 $i (i < last)$ 番目の記事を指し、 u_{last} は終端の記事 (図 24 では青窓内の右端記事) を指す。SPL(a, b) は記事 a と記事 b の SPL 値を表す。3.4.1 節の SPL 比較ではターゲット記事との SPL を見ていたが、

本節ではターゲット記事そのものを当てるためターゲット記事から一つスタート記事へさかのぼった記事である終端記事との SPL を見る。終端記事について SPL を見ようとする理由は、ターゲット記事直前であればほぼターゲット記事と同じ性質を持つ記事になっているだろうと予想できるからである。そしてテストデータ（図 19 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、テストデータの終端記事の次に来ると予想される記事名を Output する。テストデータは 1 試行当たりの記事経路長が 3～7 それぞれの中で 100 試行分をランダムに抜き出している。学習データはテストデータを除いた試行全てである。具体的なデータとして、経路長 3 の試行数は 3,688、経路長 4 の試行数は 9,488、経路長 5 の試行数は 12,591、経路長 6 の試行数は 9,231、経路長 7 の試行数は 2,263 であり、それぞれの中から 100 試行のテストデータをランダムに抜き出す。正答率の計算は、各経路長べつに 5 回学習を繰り返した時の正答率平均をとっている。また、各経路長で Wikipedia のデータ構造上全ての記事がターゲット記事になるとは限らない。それぞれの経路長でターゲット記事候補になる数が異なる。ターゲット記事候補数の平均は経路長 3 から 1337 記事、2385 記事、2755 記事、2693 記事、2265 記事である。

結果)

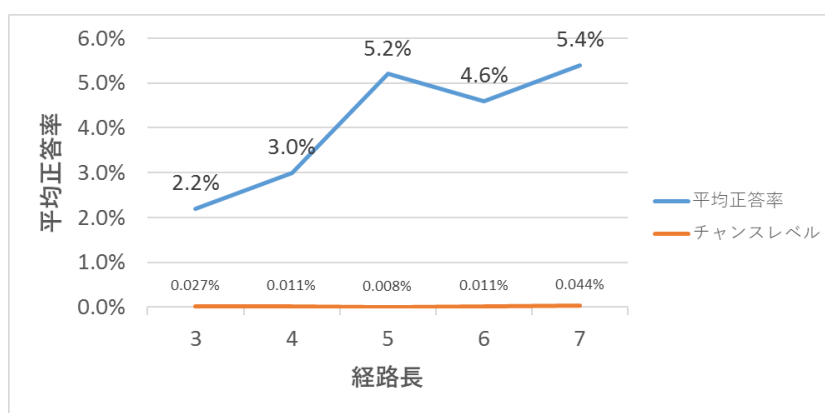


図 25 経路長別のピンポイントターゲット記事平均正答率

図 25 は経路長別にピンポイントなターゲット記事推定の平均正答率を表す。横軸は経路長、縦軸は 5 回学習とテストを繰り返した時の各正答率を平均した値である。正答率は経路長 3 から、2.2%、3.0%、5.2%、4.6%、5.4%であり、チャンスレベルは記事候補数の逆数として、0.075%、0.042%、0.036%、0.037%、0.044%であった。全体として正答率は低いものの、チャンスレベルを超える結果になった。しかし、目的推定を行うことを考えると依然として正答率は低いと考えられる。

3.4.3 ターゲット記事の絞り込み

3.4.2 ではピンポイントにターゲット記事を推定することについて考えていたが、ここでは推定された記事の中で実際のターゲット記事が何番目に確率が高く推定されているのかを確認する。これはたとえ1番目にターゲット記事が推定されていなくとも、その後続く推定確率順位で高い位置にあれば、推定された記事と似た人の行動傾向を示す記事であるかもしれない。これは、人の目的に近い位置に他の目的を知ることができているとも考えられる。そして、Wikispeediaではターゲット記事として終わりが決まっているが、現実には置き換えると終わりはなく続くものとして考えられるため、人の目的に似た他の目的を提示することでさらにその先の目的に対する行動を示せているのと同じ意味として考えられる。これは1.3.1節で述べたように、Wikispeediaが前提としているマルコフ決定過程では目的に向けて状態を遷移する様子をメタ的に見て、目的を達成したことを状態としてとらえ、目的のその先にまた目的を達成する理由となった目的があるものとして見ることができる。

使用したデータ)

3.4.2 で用いたデータと同様のデータセットを用いている。

結果)

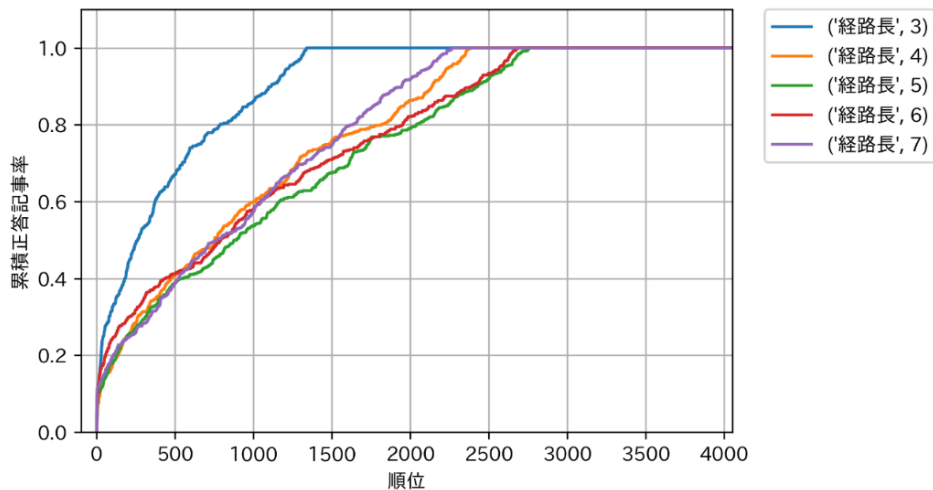


図 26 累計記事推定成功率

図 26 はある順位までに記事推定が成功した割合を示している。横軸はテストデータを与えたときに各試行に対してターゲット記事確率が高い記事から順位付けしたものを指す。縦軸は実際にターゲット記事として正解した確率を1位から累計していったものである。累計推定記事の割合が1.0になる順位が異なるのは、その経路長の試行ではターゲット記事候補になっている記事数が異なるからである。3回遷移の場合ターゲット記事になっているのは1332記事であり、4回遷移は2381記事、5回遷移は2756記事、6回遷移は2686

記事，7回遷移は2267記事であった。表の見方として，例えば経路長3（図24の青線）の場合，ターゲット候補1332位までであるうちで，500位までにターゲット記事が入っている割合は0.7に近いことが分かる。

図27より経路長3の場合はターゲット記事候補が少ないことが理由となり，経路長4以上のターゲット記事推定順位よりも早く累積正答率が高くなっている。一方で経路長4以上の正答率を見ると，順位の高いところで急激に正答率を増やし，その後は順位と相関するように正答率を伸ばしている。遷移長4以降の長さで順位が500番目ほどまでを見たとき，40%ほどの記事推定が行えていることが分かる。しかし，累計記事推定成功率が1.0になるときの順位がそれぞれの経路長でターゲット候補となる記事数と同じになることから，ターゲット記事であるのに関わらず，順位が最も低い記事として推定されていることもあると分かる。

累積正答率の傾向を詳しく見るために最もターゲット記事候補が多い経路長5のデータについて詳しく検討する。検討内容は推定された記事が最下位に近いターゲット記事が存在すること，次に順位の高い記事の性質を調べることである。仮説として，ターゲット推定は一部の記事に対して特殊化して行われているため順位が最下位に近くなるターゲット記事が存在すると考えた。

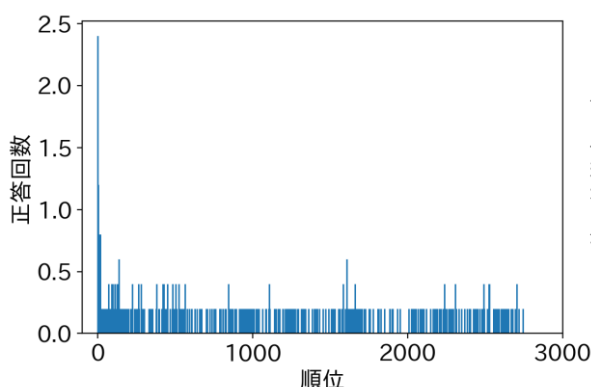


図27 記事推定成功回数（5回遷移）

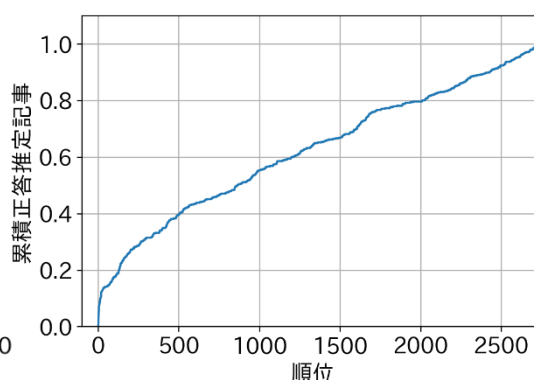


図28 累計記事推定成功率（5回遷移）

図27は経路長5のテストデータに対して各テストデータに対してターゲット記事推定された順位付けとその推定記事が正答していた回数を表す。横軸は推定された記事の順位で，縦軸はその順位の記事が実際にターゲット記事であった回数である。図28でも図27と同じく5回学習とテストを繰り返した時の平均を表している。図27から上位の順位で数回正答している他は順位に関係なく一様の正答回数になっていることが分かる。5回遷移は2756記事のターゲット記事候補があるため，最下位に近い順位のターゲット記事が存在することが確認された。図28は図27と同じく横軸に推定された順位，縦軸に累積正答率を表している。図28より経路長5のときのターゲット記事候補2762記事中，順位の上位5%

にターゲット記事がある確率は0.22, 上位10%にターゲット記事がある確率は0.30, 上位25%にターゲット記事がある確率は0.45であった。これより, 上位で正答率が高いことが分かる。しかし, 最下位近くに順位付けされるターゲット記事が存在する理由はわからない。そこで上位で記事推定が成功した記事について具体的に注目する。

表 6 学習データ記事と高順位で推定された記事の比較

	学習データ		ターゲット記事推定	
	記事名	ターゲット記事回数	記事名	推定された回数
1	Telephone	268	Telephone	33
2	Viking	212	Viking	29
3	Adolf_Hitler	69	Adolf_Hitler	16
4	Pluto	17	Banana	11
5	Banana	32	Pluto	7
6	Japan	5	Japan	3
計	-	603	-	100

表 7 は5回遷移の場合の推定記事として1位になった回数が多い6記事について, 推定回数と学習データでターゲット記事として与えられた回数をまとめている。表7より「Telephone, Viking, Adolf_Hitler, Banana, Pluto, Japan」の順番でターゲット記事推定された回数が多く, 一方で学習データとして与えられたターゲット記事の個数順は「Telephone, Viking, Adolf_Hitler, Pluto, Banana, Japan」の順番で多く見られた。これはBananaとPlutoの順番が入れ替わっただけで, その他の順番は同じであった。このことからターゲット記事として推定された回数の多い記事は, 学習データの中でターゲット記事であった数が多い記事であったと考えられる。一方で, 単純に学習データで学習された比率とターゲット記事として推定される比率が同じではなく, 表7の6記事以外にも正答している記事を高い順位で提示できていることから, 回帰分析に用いた係数にもターゲット記事推定に有効な要素があると考えられる。

3.5 分析結果まとめ

3.3 節では人の行動傾向を説明する SPL, PageRank, Outdegree の順番に分析を行った。3.3.1 節では Wikispeedia 被験者の行動は最短距離を減少し続けると仮説をたてた。検証の結果、被験者は必ずしも減少させるとは限らない遷移を行うことから仮説が支持されない結果となった。一方で最短距離はターゲット記事が近くなると急激に減少することが見られた。3.3.2 節では被験者は遷移の最初に PageRank が高くなる記事へ向かう傾向を持つと仮説を立てて検証し、全ての経路長で 1 回目の遷移で一試行中最も高い PageRank を選択する傾向にあることから仮説を支持する結果となった。3.3.3 節では被験者は一試行の最初は他記事へのリンク数を示す Outdegree を大きくする記事へ遷移すると仮説を立てて検証し、全ての経路長で 1 回目の遷移先記事の Outdegree が大きくなる記事を選択する傾向があることから、仮説を支持する結果と言える。続いて 3.4 節からはターゲット記事予測を行うために、終端記事判別、ピンポイントな記事推定、ターゲット可能性の確率順での絞り込み推定について検証した。3.4.1 節では終端記事判別を行い、結果としてチャンスレベルを超えるものの高い精度とは言えない正答率となった。一方で回帰係数を SPL 比較と PageRank のそれぞれで比較したところ、SPL はターゲット記事に近い記事の比較で値が大きく、PageRank もターゲット記事に近い記事の重みが大きいことが見られたため。3.3 節で行った人の行動傾向が利用できていることが分かった。3.4.2 節ではターゲット記事をピンポイントに 1 つの記事を当てようとした場合の正答率を出し、結果としてチャンスレベルを超えるものの精度は低くなった。これを踏まえて 3.4.3 節では 1 つの記事を推定するのではなく、ターゲット記事である確率の高い順に見た場合の正答率を出した。結果として上位 500 番目ほどの順位まで見ればテスト記事の 40%ほどで絞り込めていたことが分かったが、その絞り込みは単純に学習データとして与えられた数に起因すると仮説を立てた。具体的に遷移長 5 の試行についてターゲット予測の高い順位に出現している記事について学習データと比較したところ、たしかにターゲット記事として用いられた数にも影響されていることから仮説を支持する結果が示された。

第4章 総合議論

4.1 研究目標と研究結果から明らかになったこと

本研究で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を研究目的とし、その研究目的の達成のために具体的に次の2つの研究目標を定めていた。1つ目は行動履歴から人の目的を推定すること、2つ目は案内人が推定した目的の達成に有効な行動選択肢を提示することである。まずは1つ目の目的推定について第3章では具体的に(1)ターゲット記事に既に到着しているかどうかを判断すること、(2)未到着のときにターゲット記事のピンポイントな推定をすること、(3)未到着ターゲット記事を絞り込むこと、の順番で分析を行った。分析の内容は回帰分析から人の行動傾向を説明する各変数の正答率への寄与率の比較である。さらに、寄与率の比較を行うことで Wikispeedia 中の被験者の行動傾向から、より一般化した人の行動傾向の説明を目指し、一般化したときの目的推定アプローチについて考える。

第3章の結果から、まず終端記事がターゲット記事であるか判別を行うとき、終端記事に近い記事との SPL 比較と終端記事に近い記事の PageRank に比較的大きな重みが付いていることが分かった。しかし、判別の正答率はわずかにチャンスレベルを超えただけであり、正答率に大きく貢献しているとは言えない。次にピンポイントなターゲット記事推定についての分析ではチャンスレベルを超える正答率にはなったが、決して高いとは言えない結果となった。最後にターゲット記事の絞り込みでは、ターゲット記事である確率の高い順に見た場合の正答率を出し、全各経路長別の試行で 2000 記事中上位 500 番目までターゲットが入っていれば絞り込みが成功していると考え、テストデータの 40%ほどの試行でターゲット絞り込みを成功させている。しかし、その絞り込みは単純に学習データとして与えられた数に起因すると考え、具体的に遷移長5の試行についてターゲット予測の高い順位に出現している記事について学習データと比較したところ、たしかにターゲット記事として用いられた数にも影響されているとも考えられた。

4.2 今後の課題

本研究で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を研究目的とし、具体的に Wikispeedia 被験者の行動傾向から目的推定に有効であると考えられる要素について分析した。しかし、実施した分析は目的の推定に留まったため、有効な行動選択肢の提示まで行っていない。また、目的推定を行うための分析を行ったが各分析での正答率はチャンスレベルを超えたものの高くはないため、本研究で取り上げた行動傾向を説明する変数以外のものを探求する必要があると思われる。

第5章 結論

我々は何かしらの目的を定めて、その目的に向けてできるだけ最良の行動を計画するもの、後で振り返ったときに本当にその行動が最良であるとは限らない。そこでもし計画している行動よりもよい行動があればそれを提示してくれるリコメンドシステムが欲しいと考えた。しかし、自分で最良の行動に気づけないとき、それは単なる知識不足ではなく、自身の持つ行動選択肢が最良だと思い込んでしまうことで他の行動を検討する機会を失っていると考えられるため、そのような人に明示的に目的に向けた行動を探していると検索してもらうのは難しい。そこでリコメンドシステムが人の目的を推定することから行動選択肢の提示までを行うことを考えた。以上より本研究では、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を研究目的としていた。

研究の枠組みとして人の行動から目的を推定するために、人の行動をマルコフ決定過程モデルで抽象化することについて考えた。また、さらに簡略化した問題としてとらえるために迷路課題と目的推定課題を用いることを考えた。そして実際の人々の行動から目的を推定するために、現実世界で目的に向けて行動する人の探索行動が Wikipedia ゲームを行う被験者の探索行動に表れていると考え、Wikipedia の分析を行った。分析に用いた仮説は現実世界での人の行動仮説をたててから Wikipedia での被験者行動仮説に落とし込んだ。

分析の結果から人の行動傾向に相関する要因の一部として、Shortest Path Length, PageRank, Outdegree を特定した。目的推定の手順としてターゲット記事に到着していることの判別と記事推定、記事の絞り込みの3つについて正答率を計算したところ、それぞれチャンスレベルを超えるものの精度としては低くなった。しかしターゲット記事に到着していることの判別において、終端記事に近い SPL と PageRank の回帰係数に比較的重みがあることがみられ、人の行動傾向が正答率の向上に少しでも寄与していることが示唆された。

今後は本研究で実施した分析は目的の推定に留まったため、有効な行動選択肢の提示まで行えていない。また、目的推定分析の各分析での正答率はチャンスレベルを超えたもの的高くはないため、本研究で取り上げた行動傾向を説明する変数以外のものについて検討を行う必要がある。

参考文献

- [1] Abbeel, P., & Ng, A. Y. (2004, July). Apprenticeship learning via inverse reinforcement learning. In Proceedings of the twenty-first international conference on Machine learning (p. 1). ACM.
- [2] 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).
- [3] 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).
- [4] Melville, P., & Sindhvani, V. Recommender Systems.
- [5] 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).
- [6] Ng, A. Y., & Russell, S. J. (2000, June). Algorithms for inverse reinforcement learning. In *Icml* (pp. 663-670).
- [7] Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The PageRank citation ranking: Bringing order to the web*. Stanford InfoLab.
- [8] Richard S. Sutton & Andrew G. Barto. (2000) 『強化学習』三上貞芳, 皆川雅章訳 森北出版.
- [9] Takuya Kitazawa, 「TF-IDF で文書内の単語の重み付け」(2014). Retrieved February 7, 2019.
<https://takuti.me/note/tf-idf/>
- [10] Theory of Games and Economic Behavior. (n.d.). In Wikipedia. Retrieved February 7, 2019.
from https://en.wikipedia.org/wiki/Theory_of_Games_and_Economic_Behavior
- [11] West and Leskovec. "Human wayfinding in information networks." In WWW-12. 2012.
- [12] West, R., & Leskovec, J. (2012, May). Automatic Versus Human Navigation in Information Networks. In *ICWSM*.
- [13] West, Robert, Joelle Pineau, and Doina Precup. "Wikispeedia: An Online Game for Inferring Semantic Distances between Concepts." *IJCAI*. 2009.
- [14] 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).
- [15] 小川雄太郎(2018). 「第4回 迷路を進むエージェントの作成」,
<<https://book.mynavi.jp/manatee/detail/id=88203>> (参照 2018-10-22)
- [16] 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.

謝辞

本研究を進めるにあたり、ご助言および研究環境を提供して下さった日高昇平准教授に感謝申し上げます。日高先生には、機械系分野から認知科学の分野へ飛び込んだ自分を受け入れていただき、認知科学だけでなく研究そのものに対してもいろいろと不慣れな自分に対して粘り強い指導をして頂きました。また途中で研究方針を転向させたときにも、私の行いたい研究の意図をくみ取り支援をしていただきました。改めて感謝いたします。

また、私が苦手としていた（今もですが）プログラムや参考文献についてその読み方や使い方について教えていただいた鳥居拓馬助教授に感謝申し上げます。鳥居先生からプログラムについて教えていただいたおかげで本研究の分析は成り立っているようなものです。そして2年次になってから論文の読み方や文献の紹介をしていただき、修論執筆の時期には色々と気にかけていただいた布山さんにも感謝申し上げます。また、副テーマ指導をして頂いた橋本敬教授、副指導教員の藤波努教授にも大変お世話になりました。自分がなんとかやってこられたのも先生方のご助力のおかげです。

日高研究室第1期生として同時に入った加藤君、2年次から同じ研究室で研究をした小山君、河合君、桜井君、宮本さんにもお世話になりました。そして、研究が進まないときや、就職活動で苦勞していたときに支えてくれた友人一同にも感謝いたします。

最後に今に至る25年もの間、勉強する環境を作るために経済的・精神的支援をして頂いた家族に心からお礼申し上げます。