

地図未知な案内課題における知識の転移：知識転移の確信度の分析

Transfer of knowledge in a navigation problem without the map: analysis of confidence of knowledge transfer

鳥居拓馬[†], 日高昇平[†], 小林瞭[†]
Takuma Torii, Shohei Hidaka, Ryo Kobayashi

[†]北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology
{tak.torii, shhidaka}@jaist.ac.jp

概要

人はある課題を解くために獲得した知識を新しい課題を解くのに転用できる。こうした知識の転移を解明すべく、地図未知な意味グラフ上の案内課題を用い、この新奇課題に対して参加者が「常識」（概念の関係）を転移させるときの行動データを分析した。強化学習を用いたモデルベースの行動分析から、案内課題が終盤に向かうにつれ、参加者は転移元の「常識」が新奇課題にうまく適合したかように行動することが示唆された。

キーワード：知識の転移 (transfer of knowledge), 意味グラフ (semantic graph), 案内課題 (navigation problem), 確信度 (confidence), 強化学習 (reinforcement learning)

1. はじめに

人はある課題を解決するために獲得した知識や技能を別な新しい課題を解決するのに転用・適用できる。獲得した知識を新しい課題に適合させる能力は人がもつ問題解決の柔軟性の源のひとつといえる。通常、新奇課題に直面したとき、問題解決主体は新奇課題に関して十分な知識をもたない。そのとき、新奇課題に関する知識を探索・発見しながら、既存課題の知識を新奇課題の知識に“うまく適合させる”ことが、新奇課題の成績の向上をもたらさう。

心理学では、新奇課題を解くときに既知課題の知識を転用して成績が高まる現象は「学習の転移」と呼ばれる [1, 2]。近年の機械学習では「転移学習」と呼ばれ、特定の課題領域（課題固有の事前知識や観察データ）を超えた学習器の構築が研究されている [3]。転移学習の用語では、既知課題に関する知識を「転移の始域」と呼び、新奇課題に関する知識を「転移の終域」と呼ぶ。これらの用語を借りると、問題解決における知識の転移とは、転移の始域（既存課題の知識）と転移の終域（新奇課題の知識）の間に同型性を見いだす行為といえるだろう。その意味で、知識の転移は類推とも深く関係すると思われる。

知識の転移を経験的に研究する場合、新奇課題に対する既知課題の知識の適合度を調べるには、とくに新

奇課題の知識あるいは転移の終域が（実験者の立場から）明確に与えられていることが望ましい。そうでなければ、課題の成績の分析を超えて、新奇課題を遂行中の被験者の行動が知識の転移に誘導されたものかを分析することは難しくなる。こうした制約を部分的に充たしつつ、データ分析で扱いやすい課題として、本研究では意味グラフ上の案内課題（たとえば [4]）を扱う。

意味グラフとは、概念を頂点とし、概念間の繋がりを有向辺とするグラフをいう [5, 6]。Wikipedia 等のハイパーテキスト百科事典は、各記事を一つの概念とし、記事間のハイパーリンクを概念間の繋がりとみることで、ある種の意味グラフとみなせる [5]。本稿では、このような意味グラフを「知識」の一形式だと考える。意味グラフ上の案内課題では、問題解決主体はある始点記事から別の終点記事まで、ハイパーリンクに沿って記事を遷移して、少ない経由記事数で到達することを要求される。

典型的には、私たちはハイパーテキスト百科事典そのものを熟知することは難しく、とくに記事間の繋がりに（ハイパーリンク；以下、リンクと略記）の有無に関しては大部分が未知といえる。しかし、まったく予測できない訳ではなく、日常生活の「常識」に照らせば、記事間のリンクの有無はある程度予測できる。そこで意味グラフ（百科事典）上の案内課題では、百科事典が人間社会の概念や概念間の繋がりのある側面を表すものだとすれば、案内者（問題解決主体）は日常生活で学んだ概念や概念間の繋がりに関する「常識」と呼ばれる知識を転用して案内課題をより効率的に遂行できると考えられる。より具体的に、もし案内者の「常識」に含まれる概念間の繋がりが、百科事典に含まれる記事間（概念間）のリンクとして存在していれば、案内者はこの案内課題を少ない経由記事数で達成しう。

このように、案内課題の成績は案内者のもつ事前知識から案内する空間への知識の転移の度合いを反映するため、知識転移の可能性はその課題の成績から推定

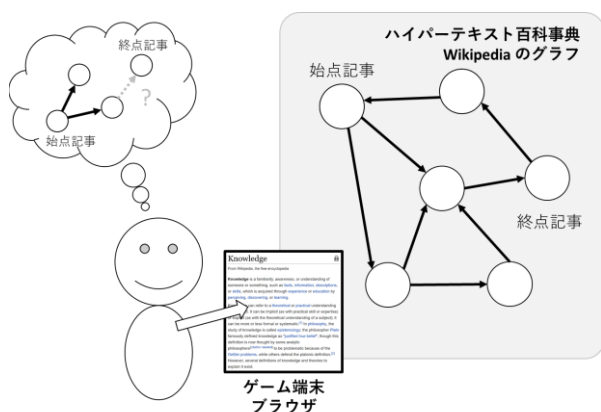


図 1. 意味グラフ上の案内課題。始点記事と終点記事が指定される。参加者（プレイヤ）のゲーム端末には各時点で訪問中の記事のテキストとハイパーリンクが表示される

できると考えられる。しかし、課題の成績のみでは、知識の転移や経路の選択を統べる問題解決の方略までは明らかとはいえない。本研究では、案内者にとって未知な構造をもつ意味グラフ上の案内課題において百科事典の記事グラフや人間の行動データを分析し、既有知識を新奇課題に転移するときの人の行動傾向を探る。

2. 意味グラフ上の案内課題と行動データ

意味グラフ上の案内課題を扱ったオンラインゲームのひとつに TheWikigame がある。TheWikigame (図 1) の各試行では、Wikipedia 内の 2 つの記事が始点および終点として指定され、始点記事から終点記事までを繋ぐ経路記事の列を、実際に Wikipedia の記事をブラウザ上で遷移しながら見つける。始点記事、経路記事の列、終点記事を並べたものを「経路」という。ゲーム開始前に始点記事と終点記事のタイトルが与えられるほか、ゲーム開始直後に始点記事のテキストおよびリンク情報が表示される（つまり、Wikipedia のある記事の全体／一部がブラウザに表示される）。表示中の記事のリンクをひとつ選択（クリック）すると、リンクで結びついた別の記事に遷移し、遷移前の記事のテキストおよびリンク情報が消え、遷移後の記事のテキストおよびリンク情報が表示される。始点から終点まで到達できた場合を「成功」と呼び、到達できずゲームを中断した場合を「失敗」と呼ぶ。各試行の成績は最終的な経路の長さ（経路長）である。ブラウザの【戻る】ボタンを使用できるが、後戻り前後の記事も経路長に加算される。Wikispeedia [4, 7] はウェブ閲覧行動の研究用に

TheWikigame を模して作られたデータ収集ソフトである。Wikispeedia で収集されたデータの一部は公開されている。本稿ではそのデータセットを分析する。

Wikispeedia データセットの基本的な統計量を述べる。Wikispeedia の記事グラフは 4,601 個の記事と 119,882 個の有向リンクをもつ。このグラフの出次数（他の記事からある記事へのリンクの数）の最小値 0, 最大値 294, 最頻値 14 である。入次数（ある記事から他の記事へのリンクの数）の最小値 0, 最大値 1551, 最頻値 0 である。全記事ペア間の最短経路長の最小値 1, 最大値 9, 最頻値 3 である。始点から終点まで到達できた参加者の行動データ（成功した経路のみ）には 51,318 本の経路が含まれる。このように、Wikispeedia の記事グラフには到達不能な記事（入次数ゼロ）や袋小路的な記事（出次数ゼロ）も僅かに含まれている。これらの特殊な記事や後戻りボタンはグラフ理論的には特殊な扱いを要するため、場合分けした上で分析することが望ましい。

そこで、本研究では Wikispeedia の記事グラフのうち最大有向強連結部分グラフ（どの記事からも任意の他の記事へのリンクの列が存在するような部分グラフ）を主な分析の対象とした。この部分グラフでは任意の記事ペアの間に少なくともひとつの有向経路が存在する。この部分グラフは 4,051 記事と 111,900 個の有向リンクを含む。グラフの統計量の最頻値は、入次数の最頻値 1 を除いて、変化していない。参加者の行動データのうち、50,792 本の経路がこの部分グラフに含まれる記事のみを用いて始点記事から終点記事まで到達した。本研究では行動データのすべての経路に対して後戻りボタンの使用等で生じる閉路を除外した“単純な”経路を抜き出し、それを分析の対象とした。

TheWikigame や Wikispeedia [4] を扱った件数はいくつもある。Wikispeedia を構築してデータを収集した [7] を含むいくつかの研究 [8, 9, 10, 11] はウェブ閲覧行動、記憶容量、情報探索の観点からデータ分析と行動モデルを研究した。本研究では、著者らは意味の認知 [12] や知識の転移といった観点から、Wikispeedia の行動データを分析する。

3. 意味グラフ上の案内行動の分析

本研究で扱う意味グラフ上の案内課題では、典型的には、案内者は Wikipedia そのもののもつリンク構造を細部まで熟知していないと考えられる。こうした案内

するグラフの全体像あるいは「地図」が未知であるとき、始点記事から終点記事までの経路を見つける課題はどれほど難しいのだろうか. 言い換えると, 案内課題の遂行において, 参加者は彼らの「常識」の知識を使っているのだろうか. これを調べるため, 人間の行動データの成績(経路長)をいくつかのベースラインと比較する. 第一のベースラインは始点記事と終点記事を結ぶすべての経路のうち最短の経路長である. 最短経路長はグラフのリンク構造に関する完全な知識(全知)を想定した場合に達成しうる成績であると同時に, 参加者が達成しうる最高の成績である. 第二のベースラインは各記事においてとりうるリンクを等確率で選択したとき, つまり一様確率のランダムウォークで終点記事に到達するまでに要する平均の経路長である. ランダムウォークの平均経路長は「常識」をもたない(無知)仮想的な主体が達成しうる平均的な成績と解釈できる.

図2は終点記事に到達できた全試行について経路長の頻度分布を計算したものである(標本数で正規化した). 人間の行動データの場合(Human), グラフ構造を既知で最短経路を選べる場合(Shortest Path), 一様確率で次の記事を選んだ場合(Random Walk)をそれぞれ図示した. 一様確率のランダムウォークの場合, 図示した範囲30ステップ時点までに終点記事に到達できる確率はほとんど0である. 図から, 人間の行動データの経路長は, グラフに関する完全な知識を想定する最短経路長ほど短くはないが, 「常識」の知識を想定しないランダムウォークの平均経路長ほど長くもない. この定性的な分析結果から, 参加者は何らかグラフ構造と 관련된「常識」をもち, それを意味グラフ上の案内課題に転用していると考えられる.

案内課題では, 始点記事と終点記事のタイトルほか, ゲーム開始直後に始点記事のテキストおよびリンク情報が与えられる. 典型的に, 始点記事と終点記事の関係は必ずしも直接的ではない. そこで, 参加者がこの案内課題を高い成績で達成するには, 始点と終点, 2つの非類似な記事(概念)を繋ぐ可能性のある経由記事をある程度まで予想した上で次の記事へのリンクを選ぶ必要がある. ひとつの仮説は, 非類似な2つの概念(始点記事と終点記事)を包括するような抽象的な概念(記事)を経由するという方略である. この仮説によれば, 相対的に具象的な概念を記す始点記事(e.g., 「イヌ」)から, 始点記事と終点記事が共に含まれる抽象的な概念を記す記事(e.g., 「動物」)を経由し, 再び相対的に具象的な概念を記す終点記事(e.g., 「バクテリア」)へ

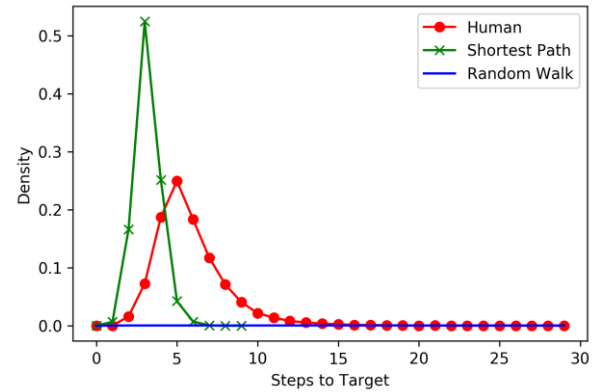


図2. 終点記事(Target)に到達できた全試行から求めた経路長の頻度分布. 人間の場合(Human), グラフ構造を既知で最短経路を選べる場合(Shortest Path), 一様確率で次の記事を選んだ場合(Random Walk)

と遷移するような経路が多く観察されるはずである. 百科事典では, 抽象的な概念を記す記事ほど下位概念を表す記事へのハイパーリンクを数多くもつ傾向にある(グラフ理論の用語では「ハブ」という). そのことから, 各記事の出次数は各記事の抽象度を表すひとつの指標となりうる.

この考えに沿って, 図3に経路長 $n = 3, 4, \dots, 9$ をもつ行動データの各時点における平均出次数(記事の概念の抽象度)を示した. 出次数による抽象度の指標では, 始点記事から経由記事そして終点記事へと遷移するにつれて, その出次数が経路の途中で高くなるような山型の形状を示している. この仮説と結果は, ウェブ閲覧行動を調査すべく [Wikispeedia](#) データセットを収集した研究者らの分析 [7] でも確認されている. さらに, 山型の傾向は人間の行動データだけでなく, グラフ構造を既知とした最短経路の場合(全知)でも見られる(図3下図). つまり, 人間の行動と最短経路はどちらも抽象概念(正確にはハブ)を経由するという仮説には合う. 注目すべき点として, 人間の行動データでは, 最初の選択(始点記事の次の記事)の時点で抽象度の平均が最大となり, そのあと平均的には単調に減少していく. 一方, 最短経路では山型の頂点は概ね経路全体の中に位置する. このことは, 人間の場合は, できるだけ早く抽象概念に到達し, その後は終点記事の概念へと着実に向かう, という方略と解釈できる.

終点記事の概念へと着実に向かう傾向は抽象度の単調減少から読みとれる. しかし, この方略では, 最短経路長(最高の成績を与える)に比べて全体的に低い成績

となっている。では、この方略は意味グラフ上の案内課題においてどのような意味があるのだろうか。

本研究の案内課題の特徴として、参加者はハイパーテキスト百科事典（意味グラフ）そのもののグラフ構造にはほとんど無知だと考えられる。それを補うために、日常生活で学んだ概念の構造すなわち「常識」を転移することで、初見の意味グラフの案内課題を効率的に遂行しているのだろうと、著者らは考えている。

喩えとして、もし私たちが地図をもたず、よく知らない町から家に帰るとしたら、できるだけ早くよく知った道にたどり着きたいと思うだろう。しかし、よく知った道に合流する経路は必ずしも最短経路とは限らない。つまり、迷うかもしれない（その知識をもつ者にとっては）近道よりも、着実に家に向かう回り道を好む人もいるだろう。これは家に着くまでの経路全体での不確かさの合計を最小化するような行動傾向と解釈できる。意味グラフ上の案内課題に戻れば、できるだけ早く抽象概念に到達し、その後は終点記事の概念へと着実に向かう、という方略・行動傾向は不確かさを減らすためだといえるだろう。言い換えると、これは知識の転移の確信度を高めるような方略・行動傾向と解釈できる。これを調べるため、次章では、人間の行動データに対して行動選択モデルのパラメータ推定を行う。

4. 案内行動データのモデリング

前章の行動データ分析から、人間の選んだ経路は、グラフ構造全知の場合（最短経路）よりも成績が低く、しかしグラフ構造無知の場合（ランダムウォーク）よりも成績が高い。この軸上では、人間の選んだ経路はこれらの間にあるものといえる。グラフ構造を既知な場合、終点記事に向かう最短経路上の記事を高く評価できる。他方、グラフ構造を無知な場合、すべての後続の記事の優劣を評価できない。こうした不確か性を伴う状況下での逐次的な意思決定に関して、各時点でのとりうる選択肢の価値づけを記述する方法のひとつに強化学習がある。

強化学習では、現在の状態 $s \in S$ 、次の状態 $s' \in S$ 、報酬関数 $R(s) \in \mathbb{R}$ に基づき、各状態の価値 $V(s)$ を見積もる。ある方略関数 $q: S \times S \rightarrow [0,1]$ のもとで、累積報酬を最大化する価値 V は以下の Bellman 最適性を満たす。関数を行列表示して

$$V = R + Q_V \gamma V$$

ここで、 $\gamma \in [0,1)$ は価値割引率、 Q_V は価値関数 V 、

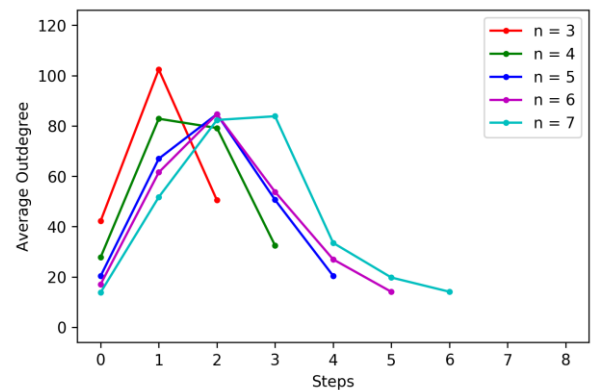
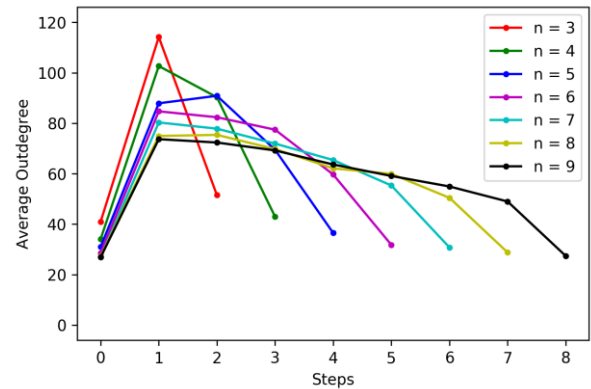


図3. 経路長 n の経路の各時点 (Steps) における平均出次数 (Average Outdegrees). 記事の出次数は記事の表す概念の一種の抽象度を表す. 上図, 人間の経路の場合. 下図, 最短経路の場合

方略関数 q , グラフ G の関数で定まる確率状態遷移行列である。本研究の用語では、状態は記事であり、方略関数はある記事にいるとき別の記事に遷移する確率を与える規則にあたる。本研究のような案内課題（迷路課題）では、終点記事である s に対して $R(s) = 1$, 他のすべての記事に対して $R(s) = 0$ と設定し、終点記事に到達することを指示する。本稿では、方略関数には softmax 関数を用いる。具体的には、ある記事 s から別の記事 s' へと遷移する確率は方略関数

$$q(s, s') \propto G(s, s') \exp[\beta V(s')]$$

とする。グラフ $G(a, b)$ は a から b へのリンクありで1を、そうでなければ0を返す。強化学習では、 β は鋭敏性パラメータと呼ばれ、 $\beta = 0$ のときは価値に依らず可能な選択肢（状態 s' ）を等確率で選ぶが、 $\beta = \infty$ のときは最大価値の選択肢（状態 s' ）の1つをほぼ確実に選ぶ。このように、 β は価値推定の確信度（あるいは不確か性）を数量化したものと解釈できる。

前章の行動データ分析から、確信度 β は時間的に変化するものと想定される。各経路 $s_0, \dots, s_t, \dots, s_{n-1}$ の

ある時点 $t \in \mathbb{N}$ で、最も尤もらしい β はすべての β のうちその記事遷移の尤度 $q(s_t, s_{t+1})$ を最大にするものである。本研究では、一定の経路長 n をもつ複数の経路標本に対してそれぞれ最も尤もらしい β を個別に求め、その平均値をみることで、経路長 n の経路の各時点でどのような確信度 β に基づく意思決定が行われたかを調べる。

強化学習では、価値関数 V はグラフ G を実質的に既知とした上で計算される。本章の分析は、あくまでも実験者の立場から、参加者の行動の背後にある価値推の確信度あるいは知識の転移の確信度を特徴づけるものである。参加者がどのような知識の転移を行っているかを直接的に調べるものではない。

図4には、各経路長 $n = 3, 4, \dots, 9$ について、計算量の理由から無作為抽出した100本の経路標本から推定した最も尤もらしい β の平均値を示した。本稿では割引率 $\gamma = 0.8$ に固定し、確信度 $\beta = 0, 1, 2, \dots, 30$ の範囲で網羅的に尤度を計算した。図から、最も尤もらしい β は始点記事の付近では小さく（相対的に $\beta = 0$ に近く）、その後、終点記事の付近に向けて次第に大きく（ $\beta = 30$ に近く）なっている。言い換えると、案内課題に取り組む参加者の行動では、各試行の経路の序盤ほど無知な場合ほどではないが確率的な選択を行っており、他方、終盤ほど全知な場合のように最適な選択を行う傾向を示している。この結果から、人間の参加者では、序盤ほど知識の転移の確信度が低く、終盤ほど知識の転移の確信度が高まる、と推察されるような行動をしていることが読み取れる。

5. 議論

本研究の意味グラフ上の案内課題（図1）では、参加者は案内課題そのもののグラフ構造（いわば地図）を知らず、その知識の欠如を補うために日常生活で培った「常識」を案内課題の問題領域に転移させるだろう、と考えた。その想定の下で行われた行動データ分析（図2）では、人間の選んだ経路はグラフ構造を全知な場合ほど高成績ではないが、まったく無知な場合ほど低成績でもないとする結果がえられた。この結果は人間が意味グラフ上の案内課題に外部の知識「常識」を持ち込んでいることを示唆すると考えられる。また、強化学習を用いたモデルベースの分析（図4）から、人間の意思決定は終盤になるほど全知な場合に近く確信度が高まっていることが示された。つまり、目的地に近づくほど

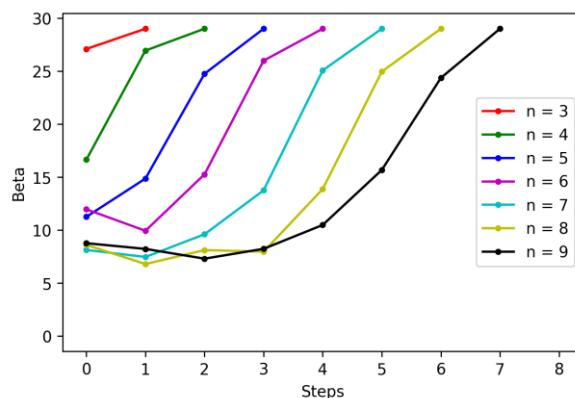


図4. 強化学習モデルから推定された価値推定の確信度を決定するパラメータ β . 始点記事での意思決定の確信度を時点0に示した。 β が大きいほど確信度が高く全知戦略（最短経路）のように振る舞う。他方、 $\beta = 0$ のときは無知戦略（ランダムウォーク）のように一様確率で選択する

「常識」をよりよく活用できる、言い換えると、知識の転移がうまくできることを示唆すると考えられる。

本研究の案内課題は、各試行の開始時に与えられる始点記事と終点記事（概念）は一般に非類似的なものである。しかし、参加者は、抽象概念を経由する方略などによって（図3）、終点記事との関連を見だしやすい記事に一端到達できれば、そのあとは比較的容易に「常識」を適用して終点記事まで到達しているように思われる。実際に TheWikigame をプレイしてみるとわかるが、終点記事に近づけば近づくほど、参加者はより正確に終点記事周辺の局所グラフ構造を予測できるだろう。その点で、各試行の終盤になるほど知識の転移の確信度が高くなるという本研究の分析結果はある程度まで納得のいくものといえるだろう。

意味グラフ上の案内行動に関していまだ未解明な点も多い。そのうちとくに興味深いものは最初の一手だと思われる。図3から、グラフ構造を既知とした最短経路の場合とは異なり、人間の行動データでは最初の一手を打ったあとは平均的には抽象度が減少し続けるような傾向を示している。このことは、最初の一手の時点で、参加者にとって終点記事との関係が比較的明らかかな、あるいはその一手の後は「常識」を転移して終点記事まで到達できそうな一手を選んでいることを示唆する。本研究では最初の一手の意思決定に関しては十分に分析できていない。図4では、人間は最初の一手から比較的高い確信度をもつことを示している。この最初の一手は囲碁や将棋などのゲームでももっとも自

由度の高くそれゆえに難しい問題といわれる。こうした最初の一手のような点に、人間の行う知識の転移のもっとも重要な側面が隠れていると考えられる。近年では、意味ネットワークの拡散反応モデル [6] が“無関係”な概念間の類似性を人間が強制判断するときの行動によく適合するという報告 [13] や意味空間のランダムウォークでよく近似できるとする報告 [14] もある。こうした研究との関連を睨みつつ、今後はベイズ統計の枠組みを用いて経路全体を包括的にモデル化するなどし、最初の一手における意思決定を分析することで知識の転移のメカニズムに迫りたい。

知識の転移は広義の類似性推論（類推や比喻等）の一種と考えられ、既存の対象から新奇な対象を理解するやり方のひとつと考えられる。将来的には、本研究で扱ったような意味グラフ上の案内課題の研究を通して、未知なる対象間の関係性を予測・発見する認知メカニズムの解明を目指していきたい。

謝辞

本研究は科学研究費補助金 JP16H05860, JP17H06713 の助成を受けて行われた。

参考文献

- [1] E. L. Thorndike and R. S. Woodworth, "The influence of improvement in one mental function upon the efficiency of other functions (I)," *Psychological Review*, vol. 8, no. 3, pp. 247-261, 1901.
- [2] 白水始, "認知科学と学習科学における知識の転移," 人工知能学会誌, vol. 27, no. 4, pp. 347-358, 2012.
- [3] L. Torrey and J. Shavlik, "Transfer learning," in *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, E. S. Olivas and et al., Eds., IGI Publishing, 2010, pp. 242-264.
- [4] R. West, J. Pineau and D. Precup, "Wikispeedia: an online game for inferring semantic distances between concepts," in *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI2009)*, 2009, pp. 1598-1603.
- [5] M. R. Quillian, "Word concepts: a theory and simulation of some basic semantic capabilities," *Behavioral Science*, vol. 12, no. 5, pp. 410-430, 1967.
- [6] A. M. Collins and E. F. Loftus, "A spreading-activation theory of semantic processing," *Psychological Review*, vol. 82, no. 6, pp. 407-428, 1975.
- [7] R. West and J. Leskovec, "Human wayfinding in information networks," in *Proceedings of the 21st International Conference on World Wide Web (WWW2012)*, 2012, pp. 619-628.
- [8] S. Aguinaga, A. Nambiar, Z. Liu and T. Wenginger, "Concept hierarchies and human navigation," in *Proceedings of the IEEE International Conference on Big Data*, 2015, pp. 38-45.
- [9] P. Singer, D. Helic, B. Taraghi and M. Strohmaier, "Detecting memory and structure in human navigation patterns using Markov chain models of varying order," in *Detecting memory and structure in human navigation patterns using Markov chain models of varying order*, vol. 9, 2014, pp. e102070-1-21.
- [10] R. Nogueira and K. Cho, "End-to-end goal-driven web navigation," in *Proceedings of the 30th Conference on Neural Information Processing Systems (NIPS 2016)*, 2016, pp. 1903-1911.
- [11] D. Helic, M. Strohmaier, M. Granitzer and R. Scherer, "Models of human navigation in information networks based on decentralized search," in *Proceedings of the 24th Conference on Hypertext and Social Media*, 2013, pp. 89-98.
- [12] T. T. Rogers and J. L. McClelland, *Semantic Cognition: A Parallel Distributed Processing Approach*, MIT Press, 2004.
- [13] S. De Deyne, D. J. Navarro, A. Perfors and G. Storms, "Structure at every scale: a semantic network account of the similarities between unrelated concepts," *Journal of Experimental Psychology*, vol. 145, no. 9, pp. 1228-1254, 2016.
- [14] J. T. Abbott, J. L. Austerweil and T. L. Griffiths, "Human memory search as a random walk in a semantic network," in *Advances in Neural Information Processing Systems (NIPS2012)*, vol. 25, 2012, pp. 3050-3058.