

潜在意味解析 (LSA) を利用した Markov 連鎖モデルによる階層メニュー探索過程の評価

北島宗雄[†] 高木英明^{††}
山本哲生^{†††}, 張 勇兵^{††}

コンピュータを利用してタスクを遂行するとき、メニューやリストのなかから目的とする項目を探し出すことが必要になることが多い。例えば、ワープロや表計算アプリケーションでは、階層メニューから適切なコマンドを選択してダイアログボックスを表示させ、文書や表の属性を編集する。WWW のなかから求めている情報を探し出す場合には、サーチエンジンが提示する候補のなかから最適なものを選択して候補ページを表示させ、タスクに関連しそうなハイパーリンクを次々と選択する。本論文では、このような項目選択タスクを、ユーザの項目選択過程の認知モデルに基づいて、(1) 項目選択をタスクとの適合度に基づいて行う、(2) 次の選択を行うときにそれまでに行ってきた項目選択の履歴を様々な仕方で考慮する、(3) ある時点で選択可能な項目の中に適当なものがなければ戻り、(4) タスクを達成できずに諦める、という 4 要素によって特徴づけられるタスクとしてとらえ、その過程を Markov 連鎖を用いてモデル化する。タスク目標が達成される確率、及び達成に至るまでの平均選択回数を算出し、それに基づいて、階層メニューやウェブサイトのリンク構造を評価する方法を提案する。

Search Process Evaluation for a Hierarchical Menu System with Markov Chains Constructed by Using Latent Semantic Analysis

MUNEO KITAJIMA,[†] HIDEAKI TAKAGI,^{††} TETSUO YAMAMOTO^{†††}
and YONGBING ZHANG^{††}

The process of successively selecting items in a hierarchical menu system to search for intended targets often constitutes an essential part in performing tasks with office application software. This paper presents Markov chain models for this process in which each menu item is a state and the degree of semantic similarity of each item to the target, given in a psychologically valid semantic space, is related to the probability of state transition for selecting the item. Then the probability and the average number of mouse clicks to reach the target are obtained by the numerical computation or the simulation of the Markov chain. We provide such evaluation for a sample hierarchical menu. Finally, a method for evaluating the structure of a menu hierarchy is proposed on the basis of this technique.

1. ま え が き

コンピュータ化されたタスクを遂行するときには、イ

ンタフェースを介してコマンドをアプリケーションに伝えることを行わなければならない。代表的なインタラクションスタイルには、メニュー、直接操作、コマンド入力があるが、WWW ブラウザ、オフィス用アプリケーション (ワープロ、表計算、作図)、ATM (現金自動預け払い機)、各種家電機器用リモコンなど広く日常生活の中で利用されているインタフェースを見れば、インタフェーススタイルの主流は、階層的に組織化されたメニューやリンク構造のなかから、項目を順次選択していくというスタイルであるといえる。

日常的な場面でタスクを遂行するために機器を利用するときには、操作の系列を記憶から呼び出してそれ

[†] 独立行政法人産業技術総合研究所
National Institute of Advanced Industrial Science and Technology

^{††} 筑波大学社会学系
Institute of Policy and Planning Sciences, University of Tsukuba

^{†††} 筑波大学社会学類
College of Policy and Planning Sciences, University of Tsukuba
現在、日立製作所
Presently with Hitachi, Ltd.

を実行するのではなく、ディスプレイ上に表示される内容を理解して、その時点で適切な項目を選択することによってタスクが遂行されることが多い。実際、メニューベースのインタラクションの場合には、たとえ繰り返し行うタスクであっても、ユーザはメニューの項目を覚えていないことが実験により確かめられている⁹⁾。このような場合、ユーザは、項目を選択するのに「ラベル追従ストラテジ」を用いることが知られている¹⁶⁾。ここで、選択の主な基準はタスクの表現と項目との意味的適合度である。したがって、項目選択を行うためのインタフェースのデザインがタスクの内容と適合していない場合には、タスクをうまく遂行できないことが予測される。

実際、このことは、メニューベースのアプリケーションやウェブのユーザビリティに関する研究のなかで確認されている。例えば、表計算アプリケーション (Excel 3.0) で新しく図を作成する場合に「ファイル」プルダウンメニューから「新規」を選択することが必要であるが、「図を作成する」というタスクゴールと選択すべき項目「ファイル」が適合していないために、タスク遂行は非常に難しかった³⁾。また、いくつかのウェブサイトを対象として、被験者にそれぞれのサイトで提供されている情報を探索させたところ、探し当てたい項目と選択すべきリンクが適合していなかったために、その成功率は 12% から 43% と非常に低かった¹⁴⁾。

本論文の目的は、階層メニューやリンク構造を介して項目選択を行いながらタスクを遂行する場合に、確実にタスクが遂行できるようなインタフェースのデザイン方法を確立するのに有用となるインタフェース評価法を提案することである。従来、メニューの評価に関する研究は階層構造に焦点を合わせて進められていて、「メニューの階層は狭く深いものよりも、広く浅いものが好ましい」というガイドラインを提案するに至っているが¹³⁾、最適な項目内容のデザインまでは考慮されていない。

本論文は次のように構成されている。まず、2 では、ユーザが項目を選択する認知過程のモデルについて述べる。次に、3 では、ユーザモデルを Markov 連鎖によって表現してインタフェースを評価する本論文の方法の特徴について述べる。4 で、項目選択過程の Markov 連鎖による定式化を示す。5 では、Markov 連鎖として推移行列を簡単に作成することができる高速評価モデルと、シミュレーションによって評価を行うモデルを導入し、結果の特徴について 5.2.2 で考察する。6 にまとめを示す。

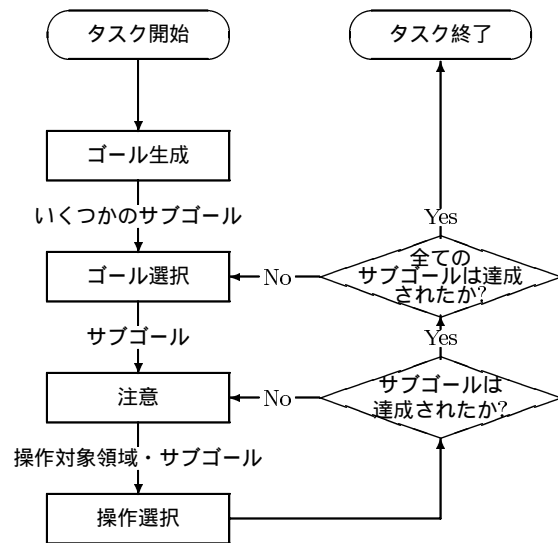


図 1 CoLiDeS モデル。
Fig.1 The CoLiDeS model.

2. 項目探索過程の認知モデル：CoLiDeS

本章では、CoLiDeS モデルの概略について述べ、次に、Markov 連鎖としてモデル化するのに重要な橋渡しとなる類似度について詳しく述べる。

2.1 CoLiDeS モデルの概略

図 1 に項目探索過程の認知モデルである CoLiDeS モデル (Comprehension-based Linked model of Deliberate Search) の概略を示す。モデルの詳細については文献 5) 6) 7) を参照されたい。CoLiDeS モデルの各過程のシミュレーションは、文章理解の認知モデルである Construction-Integration Architecture⁴⁾ に基づいて行われている。

まず、ゴール生成過程では、ユーザは最終的なゴールに到達するのに必要なサブゴールを生成する。例えば、「認知科学に関する本を閲覧する」というゴールから、「閲覧サブサイトに行く」「認知科学に関連したカテゴリーを選択する」「提示された本を閲覧する」などのサブゴールが生成される。

次のゴール選択過程では、生成されたサブゴールのなかから、現在のインタフェースディスプレイの状態に照らして最も適切なものを選択する⁷⁾。

注意過程では、インタフェースディスプレイ上に表示されている情報のうち、実行しようとしているサブゴールに関連する部分 (例えば、メニューバーやナビゲーションタブ) に注意を向ける。

操作選択過程では、注意の向いている部分から、サブゴールの表現に最も近接したオブジェクト（例えば、メニュー項目やハイパーリンク）を、意味的類似度（2.2 参照）、文字レベルでの適合度、及び今までにそのゴールの下でそのオブジェクトが選択された頻度を総合的に評価して選択する⁶⁾。

2.2 類似度の定量化法

CoLiDeS では、初めて遭遇したメニュー項目からテキストリンクを選択するとき、意味レベルでの類似度と文字レベルでの適合度を併用する。もし、現在処理しているメニュー項目中に、ゴールの表現と全く同一のラベルをもつテキストリンクが存在しないならば、意味的類似度に基づいて選択を行う。

CoLiDeS では、意味的類似度の定量化は LSA (Latent Semantic Analysis; 潜在意味解析) を用いて行われる⁸⁾。LSA は単語とそれが現れる文脈 (ドキュメント) の関係を統計的に評価する手法であり、単語および文脈は約 300 次元の意味空間内のベクトルとして表現されている。ウェブサイト <http://lsa.colorado.edu> では、米国大学生の語彙レベルに基づいて構成した意味空間をもとに、インタラクティブに、語や合成語間の類似度を得ることができる。本論文で用いた意味的類似度は、このサイトから得た数値である。

単語が複数ある場合には、合成語の意味はベクトルを合成することによって表現される。合成語間の類似度は、この意味空間を利用して、比較する 2 つのベクトルの余弦として定義される。例えば、*human computer interaction* と *software engineering* の類似度は 0.64 である。これは、これらの語が同時に現れる文脈が多い、つまり類似した語であることを反映している。一方、*parenting* という *human computer interaction* と同時に現れる文脈がないと思われる語の類似度は 0 である。このように、LSA により、意味的類似度を客観的に定量化できる。

3. Markov モデルを利用した評価の特徴

本論文では、Markov 連鎖を用いてユーザインタラクションをモデル化し、ユーザインタフェースを解析する手法を提案する。この手法は、デザインサイクルのきわめて初期の段階において、多数のインタフェースデザインを成功率や実行速度といったユーザビリティファクターの観点から極めて効率よく評価するのに有効であるという特徴をもつ¹⁵⁾。以下では、本手法と認知工学などを用いた手法との相違点を示しながら、本手法のインタフェースデザイン評価手法としての位置づけを明らかにする。

- 1) GOMS²⁾ に代表される認知モデルによれば、ほとんどすべてのデバイスとのユーザインタラクションをモデル化できる。一方、Markov モデルの方法ではデバイスの離散的な状態に基づいて分析を行うので、モデル化できるインタラクションの範囲がメニュー項目選択やボタン操作に限定される。
- 2) 認知的ウォークスルー法¹⁶⁾ 等では、ユーザの振る舞いのシミュレーションをインタラクションの個々の局面について行ってユーザビリティ評価を行う。一方、Markov モデルの方法ではすべての可能なインタラクションを対象として解析を行うことができ、結果は「平均」や「分布」によって与えられる。
- 3) Markov モデルの方法では、ゴールを達成する成功率やそれに要するステップ数がインタフェースを評価するための指標として与えられ、それらをユーザビリティファクタ¹¹⁾ の実行速度やエラー率に関連づけることができる。しかし、他のユーザビリティファクタ (学習容易性、エラーからの復旧の容易性など) や、発見的評価法¹⁰⁾ にあげられている項目については、直接的に有効な指標は提供できない。総合的な評価を行うためには、それらと組み合わせる必要がある。
- 4) 複雑な機能を有したデバイスのインタフェース評価をユーザテストによって行う場合には、極めて少ない数のインタフェースデザインしか評価の対象とすることができない。一方、Markov モデルの方法では、ユーザモデルを用いて統計的にユーザがシステムの「全体」とどのようにインタラクションする可能性があるかについて調べることができる。種々のユーザモデルを用いることにより、数多くのインタフェースデザインの評価を、半ば自動的に、効率的に、しかも、実際のユーザよりも完璧に行うことができる。
- 5) ユーザテストやインスペクション法を行うには、表示位置などのユーザインタフェースデザインが詳細にわかっている必要があるため、これらをインタフェースデザインの初期に利用することはできない¹¹⁾。一方、Markov モデルの方法は、インタフェースデザインのきわめて初期の段階 (インタフェースの機能仕様の決定前) で適用できる。ただし、デバイスの離散的な状態は記述されなければならない (たとえば、すべてのメニュー項目とそれらのリンク構造)。

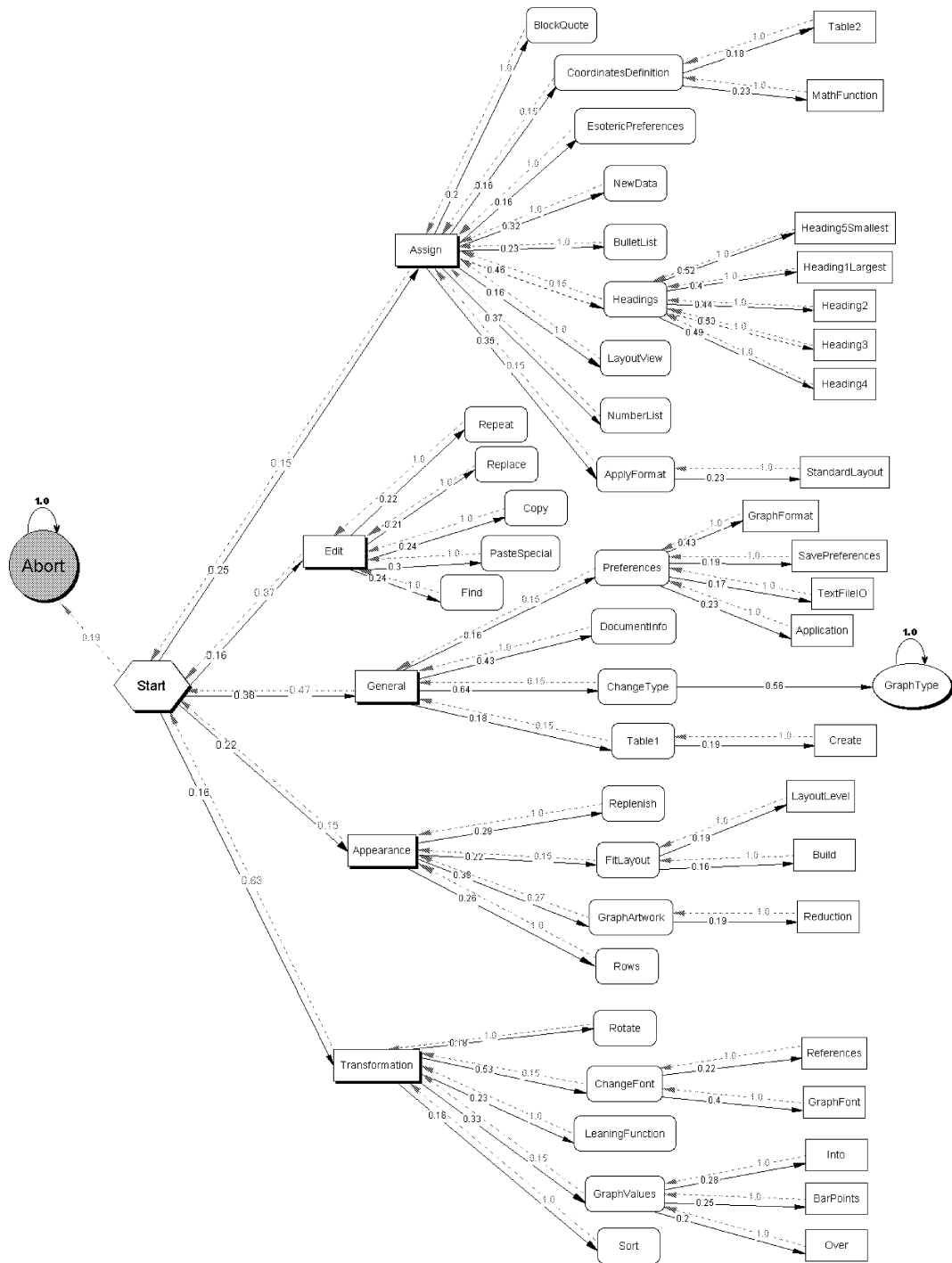


図 2 階層メニューと類似度 (履歴を利用しないモデル).
 Fig.2 A hierarchical menu with similarity indices (a memoryless model).

4. 項目選択過程の Markov モデル

本章では、メニューバー、プルダウンメニュー、サブメニューの3階層からなる階層メニューを想定し、そ

こから目的とする項目を選択する場合について、その成功率とそこに至る項目選択回数の平均値を求める方法について説明する。まず、例として取り上げる階層メニューについて説明し、次に、ユーザの項目選択過

程を Markov 連鎖としてモデル化する方法について述べる。

4.1 階層メニューの表現と類似度の付与

図 2 に本論文で用いる階層メニューの一部を示す。後で説明する理由により選択される可能性のない項目は実際の階層メニューから除かれている。ユーザは、ラベル追従戦略¹⁶⁾を用いて階層メニュー内の項目を次のように選択すると仮定した。

- 1) メニューバー項目を選択してプルダウンメニューを表示させる。あるいは、タスクを終了する。
- 2) プルダウンメニューのなかからメニュー項目を選択し、ダイアログボックスまたはサブメニューを表示させる。あるいは、メニュー項目を選択せずに上の階層に戻る。
- 3) サブメニューが表示された場合にはそのなかからメニュー項目を選択し、ダイアログボックスを表示させる。あるいは、メニュー項目を選択せずに上の階層に戻る。
- 4) 正しいダイアログボックスが表示されればタスクを終了する。そうでない場合には、上の階層に戻る。

なお、現実のダイアログボックスの場合には、キャンセルボタンなどが備えられていて、ユーザは誤った選択を行った場合、それを押して一番最初に戻ることが考えられる。インタフェースの具体的な機能仕様も考慮して認知ウォークスルー法¹⁶⁾などによりメニュー項目の評価を行う場合には、それを考慮したユーザのメニュー項目選択過程のモデルを仮定する必要があるが、本論文では、3 で述べたように、メニュー項目とリンク構造が決まった時点での評価を想定しているので、キャンセルボタンなどのインタフェースの機能仕様を切り離してユーザのメニュー項目選択過程をモデル化している。

メニューバー項目は、Edit, General, Transformation, Assign, Appearance の 5 項目があり、メニューバー項目が選択されると、プルダウンメニューが表示される。さらに、プルダウンメニューを選択した場合には、サブメニューが表示される場合と、ダイアログボックスが表示されるなどコマンドが発効される場合がある。サブメニューを選択した場合には、ダイアログボックスが表示される。

図 2 で実線で示した矢印についている数値は、ユーザが、“change the graph type to column”（グラフのタイプをコラムにする）というタスクを行おうとしているとき、そのタスクと、各メニュー項目の間の意味的類似度である。これらは、2.2 で述べたように LSA を用いて定量化される。なお、本論文では、意味空間

として米国の大学生の語彙レベルを表現する意味空間を用いている。このタスクにおいて、ターゲットまでの最短パスは、Start → General → Change Type → Graph Type である。Graph Type を選択することにより、タスクに適合したダイアログボックスが表示されるので、成功したことがわかる。

しかし、実際には、正しくないダイアログボックスが表示されたり、選択肢の中にメニュー項目がないために上の階層に戻ることがある。図 2 で破線で示した矢印についている数値は、上の階層に戻る場合のリンクのウェイトを示している。正解でない端末ノードについては、必ず上の階層に戻るため、リンクウェイトは 1.0 である。正解である端末ノードと、タスク放棄を表す Abort ノードには該当するリンクは存在しない。その他のノードについては、図 3 に示した手続きで、戻りリンクのウェイトが決定される。すなわち、ユーザはラベル追従戦略を採用して項目選択を行うと仮定しているため、類似度が小さい項目が選択される可能性はない。そこで、ある閾値を定め（ここでは、0.15 とする）、タスクゴールとの類似度が閾値以下のノードについては選択される可能性がないとして、その類似度を戻りリンクのウェイトに繰り入れる。例えば、図 3 で、メニューバー項目 Edit のプルダウンメニューには 9 個の項目が存在するが、そのうち 4 項目の類似度は閾値以下である。そこで、Edit からの戻りリンクのウェイトは、それらを加えて、0.37 となる。また、Start から出るパスにおいて、そのような状態がある場合は、Abort へのリンクのウェイトに加える。なお、閾値以下の類似度を加えても閾値に達しない場合や、閾値以下の類似度を持つ選択項目が存在しない場合には、戻りリンクのウェイトとして閾値の値を設定した。

このようにして、類似度をもとにしてリンクにウェイトを付与し、その大きさに応じた確率で状態推移を起こさせることにより、Start から始めて、ターゲットもしくは Abort に到達するまで、有限時間でこのタスクを終了することができる。

なお、本論文では、タスクとメニュー項目の類似度を、簡単のために、与えられたタスク記述 “change the graph type to column” とメニュー項目の表現、たとえば “Graph Type” を用いて LSA により得た。しかし、実際のユーザは、タスクやメニュー項目が与えられるとそれらに関連した知識を活性化して、それらを含めて類似性の評価を行うと考えられるので、本論文での推移確率が現実のものとは異なっている恐れがある。本論文では、詳細に述べていないが、2 で説明した項目

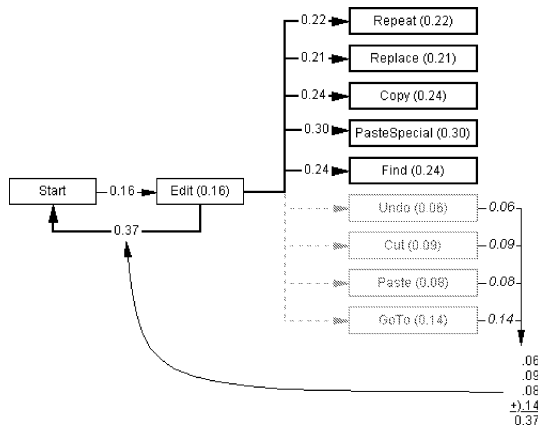


図3 「戻る」に対するウェイトの付与。

Fig.3 Assignment of weight for "backward" transition.

探索過程の認知モデル⁵⁾⁶⁾⁷⁾の操作選択過程には、タスク記述やメニュー項目の表現を選択に先立って精緻化する過程が含まれている。その過程を考慮することにより、よりユーザの判断に近い類似度が得られると考えられる。実際、ユーザが探している項目を探することができるかどうかという観点からウェブユーザビリティを評価するウェブ認知ウォークスルー (Cognitive Walkthrough for the Web)¹⁾¹⁸⁾では、精緻化された見出しやリンクを用いて類似度を求めている。そして、被験者実験により、探している項目と正しい項目の間の類似度が小さい場合に、その項目を選択することが難しくなり成功率が低くなることが確認されている¹⁾。精緻化過程を考慮して求めた類似度を用いることにより現実と適合した推移確率が得られることが期待されるが、本論文では、提案する評価法の記述の簡単化のためにその過程を考慮していない。

4.2 Markov 連鎖による定式化

特定の階層メニューにおける項目選択過程は、選択可能な項目が並んだメニューを状態とし、クリックを時刻パラメタとすることにより、離散的状態空間と離散的時刻をもつ Markov 連鎖でモデル化することができる。メニューの種類は $r (< \infty)$ 個であると仮定し、状態空間を $S = \{1, 2, \dots, r\}$ とする。 n 回目のクリックで現れるメニューを X_n とする。このとき、状態空間 S をもつ Markov 連鎖 $\{X_n; n = 0, 1, 2, \dots\}$ の推移確率

$P_{ij} := P\{X_{n+1} = j | X_n = i\}$ $i, j \in S$ が与えられているとする。

Markov 連鎖の理論によれば、非再帰の状態 i から再帰の状態 j への初到達ステップ数 T_{ij} が有限である

確率

$$f_{ij} := P\{T_{ij} < \infty\}$$

と、 T_{ij} が有限であるときの平均初到達ステップ数

$$\mu_{ij} := E[T_{ij}; T_{ij} < \infty]$$

は、固定した j に対し、次の線形連立方程式を満たす¹²⁾。

$$f_{ij} = \sum_{k \in C(j)} P_{ik} + \sum_{k \in T} P_{ik} f_{kj} \quad i \in T \quad (1)$$

$$\mu_{ij} = f_{ij} + \sum_{k \neq j} P_{ik} \mu_{kj} \quad i \in T \quad (2)$$

ここで、 $C(j)$ は再帰の状態 j が属する相互到達可能な状態の同値類を表し、 T は Markov 連鎖の非再帰の状態の集合を表すものとする。従って、 $\{P_{ij}\}$ が与えられていれば、固定した再帰の状態 j に対し、これらの方程式を解いて $\{f_{ij}; i \in T\}$ と $\{\mu_{ij}; i \in T\}$ を求めることができる。

4.3 推移確率行列の作成

状態推移確率行列が与えられた Markov 連鎖を用いると、タスクパフォーマンス指標を解析的または数値的に求めることができる。本節では、図2に示した状態推移について、推移確率行列を作成する方法を述べる。

離散的時刻をもつ Markov 連鎖を考えると、推移確率行列は、状態ごとに適当に番号を付け、状態 i から状態 j に行く推移確率を (i, j) 成分とする。4.1に示した方法により計算した類似度を各行和が1になるように比例配分し、推移確率に変換する。56個の全状態に対し、Start から端末ノードへ順に状態の番号をつける。便宜上、ターゲット ("Graph Type") を状態 55 とし、Abort を状態 56 とすると、推移確率行列は以下のように表される。

$$(P_{ij}) = \begin{pmatrix} P_{11} & \dots & P_{1,54} & P_{1,55} & P_{1,56} \\ P_{21} & \dots & P_{2,54} & P_{2,55} & P_{2,56} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ P_{54,1} & \dots & P_{54,54} & P_{54,55} & P_{54,56} \\ \hline 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 0 & 0 & 1 \end{pmatrix}$$

この Markov 連鎖では、状態 1~54 が非再帰の状態であり、状態 55 と 56 が再帰的 (吸収) 状態である。ターゲットに至るアクセスだけについての平均クリック数は

$$\frac{\mu_{1,55}/f_{1,55}}{\text{ターゲットに至るアクセスのクリックの合計数}} \\ \text{ターゲットに至るようなアクセスの数}$$

で与えられる。同様に、Abort に至るアクセスの平均クリック数は、 $\mu_{1,56}/f_{1,56}$ で与えられる。

5. 階層メニューの評価

本章では、ユーザの項目選択過程モデルをあらたに 4 種類導入し、具体的な数値をあてはめて項目選択過程を評価する。評価結果は成功率、失敗率、及び平均クリック数によって与えられるが、それらを各ユーザモデルを想定した場合のタスク実行速度に関連づけることが可能である。本論文のユーザモデルは、GOMS モデル²⁾のようにタスク分析に基づいていないので、ユーザの行動を正確にシミュレートできるわけではない。しかし、評価結果に基づき、いくつかのインタフェースデザイン候補の相対的な比較を行うことは可能である¹⁵⁾。たとえば、ユーザが目的を達成するまでの過程を Markov 連鎖でモデル化して e-コマースサイトの web デザインを評価する試みが報告されている¹⁷⁾。

5.1 5 種類のユーザのモデル

5.1.1 ユーザモデル M1

ユーザモデル M1 に対する状態推移図は図 2 に示したものである。このモデルにより、ターゲットに到達する確率は正確に計算できるが、そのときのクリック数の評価は現実的ではない。

5.1.2 ユーザモデル M2

もともと Markov 連鎖では、過去と現在のすべての履歴の条件の下での将来の状態確率は現在の状態のみに依存し、過去の履歴とは独立であるという Markov 性を持っている。しかし、実際の項目選択過程では、一度訪れた状態にターゲットが存在しなかった場合、もう一度その状態を訪れることは考えにくい。つまり過去の履歴に少なからず影響を受ける。図 4 に示したのは、あるメニュー項目を選択したあとで選択元に戻ってきた場合に、直前に選択した項目は選択しないというユーザモデル M2 である。例えば、メニューバーから General を選択して、表示されたプルダウンメニューから Preferences を選択する (Start → General → Preferences)。そこで、Preferences のサブメニューを選択せずに「戻る」を選択したとする。このとき、戻り先は図に示したように “General ← Preferences” と表示されたノードである。このノードには Preferences へのリンクはなく、General のその他のプルダウンメニューへのリンクのみが存在する。この時点で、どのプ

ルダウンメニューも選択せずに戻る場合には、“Start ← General” と表示されたノードに戻る。この状態では、General へのリンクは切られている。

5.1.3 ユーザモデル M3

M2 ではメニュー階層の間を何回も往復することができ、ターゲットまたは Abort にたどり着くまでに非常に多くの推移を行うことになる。時間が無限にある場合を除いて、現実的なメニュー探索方式とは考えられない。図 5 は、図 4 を若干修正して、深さ優先の探索方式を表現するとともに「戻り」に制約を加えて、探索が速く終了するようにしたユーザモデル M3 を示す。図 4 との違いは、2 段階上の階層には戻ることができず、その場合には Abort する点である。例えば、“General ← Preferences” の状態で「戻る」が選択されたとき、“Start ← General” ではなく、Abort に推移する。

5.1.4 ユーザモデル M4

M1 ~ M3 については推移確率行列を作成することができるという利点があるが、実際のユーザの項目選択過程のモデルとしてはあまり現実的ではない。現実には、一度選択したリンクは再度選択しないというモデルが適当であろう。ユーザモデル M4 では、図 2 の状態推移において、一度訪れたノードは 2 度と訪れないと仮定する。例えば、Start → Appearance と推移し、Appearance プルダウンメニューから何も選択せずに Start に戻ったとき、Start から Appearance へのリンクは完全に断ち切られる。したがって、この時点で、Appearance の下の階層へは到達できないことになる。このように、戻ることによって、まだ訪れていないにもかかわらず永久に到達できない部分ができてしまうことから、このモデルは、慎重に「戻る」を選択するユーザのモデルを表現しているといえる。これは、ウェブなど、一度訪れたことが明示的に表示される場合に、ユーザが取る行動のモデルとして適当であろう。

5.1.5 ユーザモデル M5

M4 を変形したユーザモデル M5 により、端末ノードのみ再度訪れないユーザを表現する。図 2 に示したように、端末ノードが特定のメニュー階層の深さに対応しているわけではない。どれが正しくない端末ノードであったかを完全に覚えているという意味で、これは記憶のよいユーザのモデルとなっている。これは、オフィスアプリケーションなどで誤ったコマンドを再発効したり誤ったダイアログボックスを再表示させずに項目選択を行うユーザのモデルとして適当であろう。

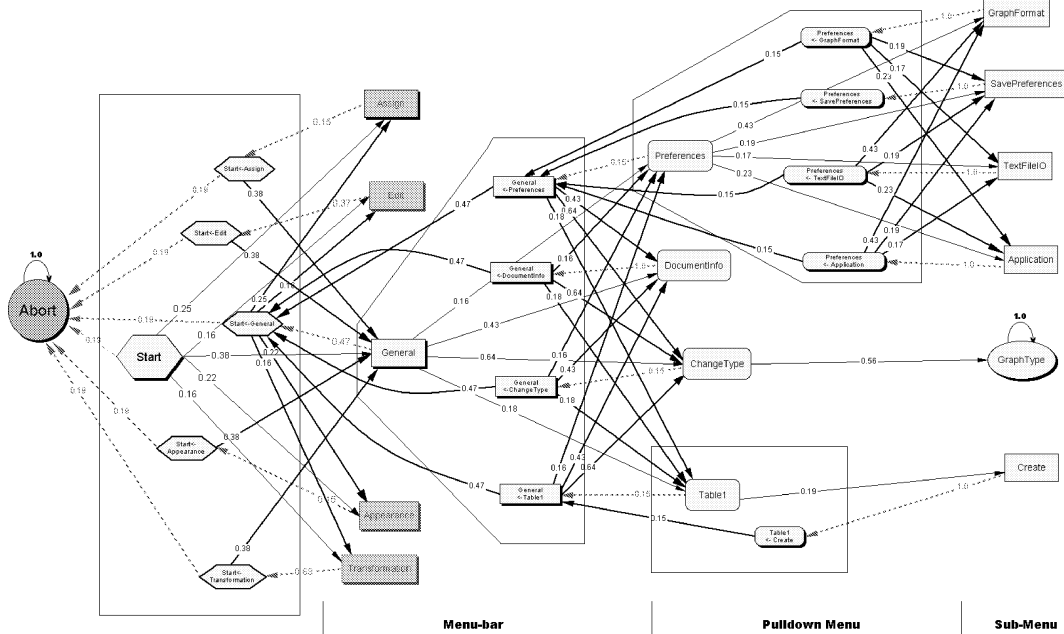


図 4 ユーザモデル M2 .

Fig.4 A Markov model of users - M2.

5.2 階層メニューの評価

5.2.1 ユーザモデルによる評価結果

表 1 に Start からターゲットまたは Abort への初到達確率 $f_{1,55}$, $f_{1,56}$ (成功率, 失敗率) と平均クリック数 $\mu_{1,55}/f_{1,55}$, $\mu_{1,56}/f_{1,56}$ を示した. ユーザのモデル M1 ~ M3 については, 推移確率行列を作成することができるので, 4.3 で説明した方法によってこれらを算出することができる. モデル M4 と M5 については, 項目選択シミュレーションを 10,000 回行って, 成功率, 失敗率, 及び平均クリック数を求めた. 表 1 には 95 % 信頼区間も示してある.

5.2.2 考 察

表 1 に示された評価結果に基づいて, 履歴を考慮することによる影響について考察する. M1 と M2 の計算結果を比較すると, M2 の平均クリック数は成功, 失敗いずれの場合も大幅に減少していることがわかる. これは, 「戻る」が選択された後の選択項目数が履歴を考慮しない場合 (M1) よりも 1 つ少ないためである. 成功率については, 履歴を考慮することによって若干減少しているが, これも, 原因は同じである. 例えば, Start \rightarrow General \rightarrow Start と推移した場合, 履歴を考慮しない場合 (図 2) に Abort に推移する確率は $0.19/(0.19 + 0.25 + 0.16 + 0.38 + 0.22 + 0.16) = 0.088$ であるが, 履歴を考慮する場合 (図 4) には, 「Start \leftarrow General」の状態から Abort に推移するの

で, $0.19/(0.19 + 0.25 + 0.16 + 0.22 + 0.16) = 0.107$ となる. これに反して, ターゲットへの推移確率は変わらない. このように, Abort への推移が可能なる場合について推移確率が大きくなるので, 成功率が減少する.

M3 において成功する場合を見ると, 平均クリック数が 4.1 であり, 最短パス長 (Start \rightarrow General \rightarrow Change Type \rightarrow Graph Type) に非常に近い値になっている. 一方, 失敗する場合の平均クリック数は 10.8 であり, 何回かの試行錯誤の後, Abort に推移することがわかる. M3 では, 深さ優先で階層をまたがって戻ることができない場合の特徴が捉えられていると考えられる. M4 も, M3 と同様な結果となっているが, M3 では許されなかった階層をまたがる戻りが許されているために, 成功率と平均クリック数が増大している.

M5 は正しくない端末ノードへの再訪のみを禁止しているため, 当然, 成功率, クリック数とも M4 よりも大きくなっている. 状態推移図は M1 とほぼ同じであるが, 履歴を考慮することによって, 成功率は若干向上し, 平均クリック数は大幅に減少している. したがって, M1 による成功率は M5 の成功率に近い値を提供すると考えられる.

5.3 階層メニュー評価法の提案

以上のようにして, 階層メニュー構造とユーザのターゲットが与えられたときに, 成功率と平均クリック数

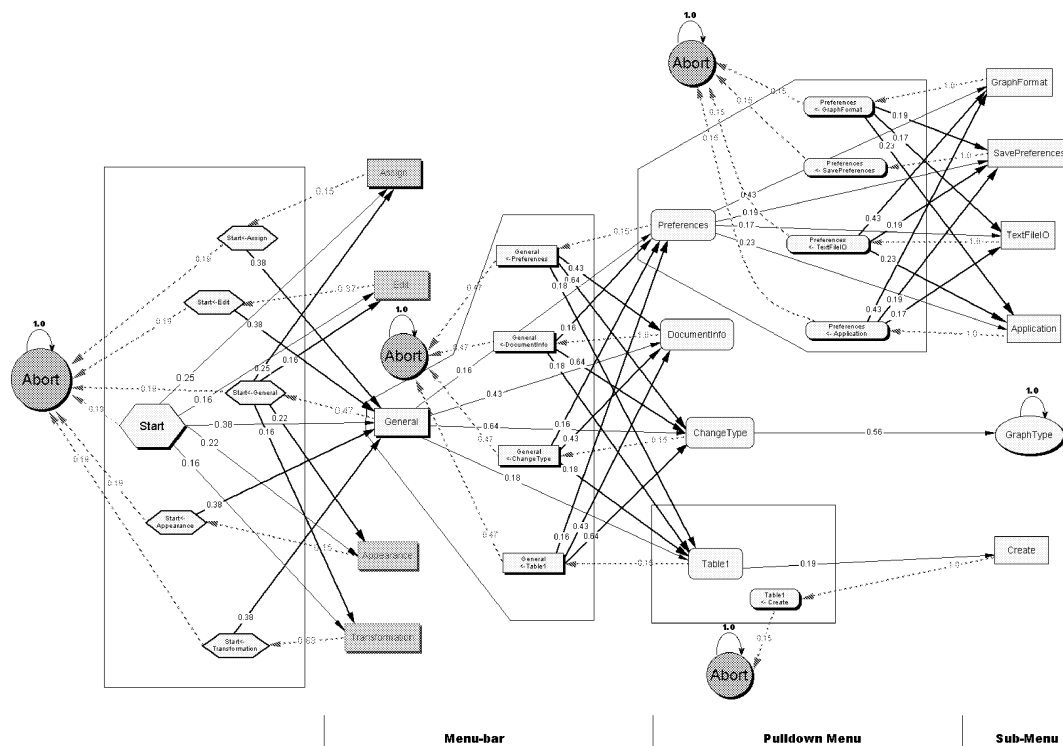


図 5 ユーザモデル M3 .

Fig.5 A Markov model of users - M3.

表 1 成功率, 失敗率と平均クリック数 .

Table 1 Success and failure probabilities and average number of clicks.

モデル名	成功する場合				失敗する場合		
	成功率	(95%信頼区間)	平均 クリック数	(95%信頼区間)	失敗率	平均 クリック数	(95%信頼区間)
M1	.509	—	131.2	—	.491	124.6	—
M2	.451	—	85.9	—	.549	81.5	—
M3	.124	—	4.1	—	.876	10.8	—
M4	.304	(0.294,0.313)	16.3	(15.8,16.9)	.696	24.3	(23.9,24.8)
M5	.520	(0.510,0.530)	32.3	(31.5,33.1)	.480	27.3	(26.5,28.1)

を 5 種類のユーザモデルを用いて求めることができる .
 すでにモデルを導入する際に説明したように, M1~M3 は実際のユーザの項目選択過程のモデルとしては精度はよくないが, 数値計算により評価結果を短時間に得ることができるので, 高速に階層メニューの相対的な評価を行うことができる . 一方, M4 と M5 は実際のユーザに近い振る舞いはするものの, 評価結果を得るのに時間がかかる . この特徴を利用して, 以下に提案する方法によって, 階層メニューの評価をインタフェースデザインの初期の段階で効率的に行うことができる .

- 1) 想定されるターゲットを列挙する (端末ノードの表現で代用できる) .

- 2) 候補となる階層メニューをいくつかデザインする .
- 3) すべてのターゲットの集合と階層メニュー候補の組合せについて, 高速評価により成功率と平均クリック数を求める . 評価の高かった階層メニューについてシミュレーションによる評価を行う . このとき, 階層メニューがウェブタイプなのか, オフィスアプリケーションタイプなのかによって, M4 と M5 を使い分ける .

6. む す び

本論文では, IT によってもたらされる豊かな情報環境とのインタフェースに欠かすことのできない階層メニューの評価法について述べた . ユーザが項目選択を

行う過程の認知モデルに基づいて、階層メニューからターゲット項目を選択する過程を Markov 連鎖としてモデル化し、ユーザの項目選択過程に応じて、5 種類のモデルを導入した。このうち、3 種類のモデルは、推移確率行列を作成することができることから、高速に成功率と平均クリック数を求めることができる。残りの 2 種類のモデルは、ウェブあるいはオフィスアプリケーションにおいてユーザが項目選択をする過程の特徴を考慮したモデルになっていることから、ユーザがタスクを遂行する環境の特殊性を考慮したより精度の高い評価を行える。本論文では、階層メニューのひとつの例について、これらのモデルによって実際に成功率と平均クリック数が求められ得ることを示し、さらに、インタフェースデザインの初期の段階に適用できる階層メニューの評価法を提案した。ここで提案した階層メニュー評価法の実装については、現在作業を進めているところであり、多数の実例を用いた有効性の検証は今後の課題である。

参 考 文 献

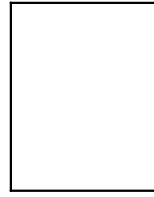
- 1) Blackmon, M. H., Polson, P. G., Kitajima, M. and Lewis, C.: Cognitive Walkthrough for the Web, *Proceedings of the SIGCHI conference on human factors in computing systems*, New York, ACM, ACM Press, pp. 463-470 (2002).
- 2) Card, S. K., Moran, T. P. and Newell, A.: *The Psychology of Human-Computer Interaction*, Lawrence Erlbaum Associates, Hillsdale NJ (1983).
- 3) Franzke, M.: Turning research into practice: Characteristics of display-based interaction, *CHI '95 Proceedings*, New York, ACM, ACM Press, pp. 421-428 (1995).
- 4) Kintsch, W.: *Comprehension: A Paradigm for Cognition*, Cambridge University Press, Cambridge (1998).
- 5) Kitajima, M., Blackmon, M. H. and Polson, P. G.: A comprehension-based model of Web navigation and its application to Web usability analysis, *People and Computers XIV - Usability or Else!* (McDonald, S., Waern, Y. and Cockton, G.(eds.)), London, Springer, pp. 357-373 (2000).
- 6) Kitajima, M. and Polson, P. G.: A comprehension-based model of correct performance and errors in skilled display-based, human-computer interaction, *International Journal of Human-Computer Studies*, Vol. 43, pp. 65-99 (1995).
- 7) Kitajima, M. and Polson, P. G.: A comprehension-based model of exploration, *Human-Computer Interaction*, Vol. 12, pp. 345-389 (1997).
- 8) Landauer, T. K. and Dumais, S. T.: A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge, *Psychological Review*, Vol. 104, pp. 211-240 (1997).
- 9) Mayes, T. J., Draper, S. W., McGregor, A. M. and Oatley, K.: Information flow in a user interface: The effect of experience and context on the recall of MacWrite screens, *Proceedings of the HCI'88 Conference on People and Computers IV*, Cambridge, Cambridge University Press, pp. 275-289 (1988).
- 10) Molich, R. and Nielsen, J.: Improving a human-computer dialogue, *Communications of the ACM*, Vol. 33, pp. 338-348 (1990).
- 11) Newman, W. and Lamming, M.: *Interactive System Design*, Addison-Wesley, Reading MA (1995). (邦訳: インタラクティブシステムデザイン (北島宗雄, 監訳), ピアソン (1999)).
- 12) Parzen, E.: *Stochastic Processes*, Holden-Day, San Francisco (1962).
- 13) Shneiderman, B.: *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, Addison-Wesley, Reading MA, 2nd edition (1992). (邦訳: ユーザーインタフェースの設計: やさしい対話型システムへの指針 (東基衛・井関 治, 監訳), 日経 BP 社 (1993)).
- 14) Spool, J. M., Scanlon, T., Schroeder, W., Snyder, C. and DeAngelo, T.: *Web Site Usability: A Designer's Guide*, Morgan Kaufmann, San Francisco (1999). (邦訳: Web サイトユーザビリティ入門~サイトの「使いやすさ」を考える~ (篠原 稔和, 監訳), トッパン (2000)).
- 15) Thimbleby, H., Cairns, P. and Jones, M.: Usability analysis with Markov models, *ACM Transactions on Computer-Human Interaction*, Vol. 8, pp. 99-132 (2001).
- 16) Wharton, C., Rieman, J., Lewis, C. and Polson, P.: The cognitive walkthrough method: A practitioner's guide, *Usability inspection methods* (Nielsen, J. and Mack, R. L.(eds.)), John Wiley and Sons, New York, pp. 105-140 (1994).
- 17) 狩谷典之, 北島宗雄, 高木英明, 張勇兵: Markov モデルを用いた e-コマースサイトの web デザイン評価, 電子情報通信学会誌 B, pp. 1809-1812 (2002).
- 18) 北島宗雄: ウェブ認知ウォークスルーによるウェブサイトユーザビリティの評価, ファジィ学会誌 (印刷中), Vol. 14, pp. XXX-XXX (2002).
(平成?年?月?日受付)
(平成?年?月?日採録)



北島 宗雄

独立行政法人産業技術総合研究所
人間福祉医工学研究部門認知的インタ
フェイスグループ研究グループ長。
博士(工学)。昭 53 東工大・理・物
理卒。昭 55 同大学院修士課程了。

同年通商産業省工業技術院製品科学研究所入所。生命
工学工業技術研究所主任研究官を経て平 13 より現職。
ヒューマンコンピュータインタラクションにおける認
知モデリングの研究に従事。



山本 哲生

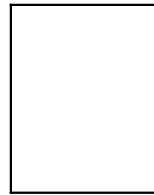
平 13 筑波大・第三学群・社会工学
類卒。同年(株)日立製作所入社。



高木 英明

筑波大学社会工学系教授。Ph.D.
昭 47 東大・理・物理卒。昭 49 同大
大学院修士課程了。同年日本アイ
ビー・エム(株)入社。昭 58 カリフォ
ルニア大 Ph.D.(計算機科学)。平 5

IBM 東京基礎研究所主幹研究部員。同年筑波大・社会
工学系教授。情報通信ネットワークと確率モデルの研究
に従事。



張 勇兵(正会員)

筑波大学社会工学系助教授。博士
(工学)。昭 59 職業能力開発大・電
気科卒。平 4 電通大学院博士課程
了。同年同大情報工学科助手。筑波
大・社会工学系講師を経て平 12 より

現職。分散コンピュータシステム、通信ネットワーク
などの研究に従事。