論 文

# 推移確率に基づく対話型事象検索システム*

磯本征雄** 山県敬一***

**Abstract**

This paper deals with the searching mechanism of the event in accordance with the user's aim in a large event set. The pseudo-ordered event set with the probabilistic weights of the relation is introduced, and is formulated as a Markov chain model mathematically. Using the consequent from the probability theory, the event sequence whose events are intensively associated with each other, is easily detected by the computer.

In order to deal with the probabilistic model, the interactive processes are essential. The directed graph, which represents the Markov chain, is always updated through man-machine interactions. Moreover, both the adjacency matrix and the transition probability matrix are modified by the user's response. Finally, user can find out his goal event efficiently. The debugging process of FORTRAN programs is also discussed as an application of this event search system.

## 1. まえがき

ある有限な事象集合において，事象相互の間に2項関係が定義されているとき，この関係を有向グラフで表すことができる[1]．さらに，この有向辺に関係の強さを表す確率論的重みが関係を満たす事象の組ごとに独立に与えられるものとすれば，正規化することによりこの有向グラフは一重マルコフ連鎖グラフとなる．マルコフ連鎖[2]でモデル化できる対象は多く存在するので，これに基づく事象検索問題はいろいろな応用を持つと考えられる[3),4)]．この場合の利点として，確率論的重みに依存する検索が確率論の手法を用いて行列演算で実行できることと，グラフが単純であって特定の2事象を結ぶ有向辺の数は 1 か 0 であり，隣接行列[5)]がビット・マトリックスで表現できることの2点が挙げられる．これらの利点を活かして，決定論的モデルによるグラフ上の探索問題とは別の観点から，事象検索を扱う試みについてはすでに論じた[6)]．

本論文では文献 6) の考え方をさらに発展させ，より大きなグラフに対しても効率よく検索手順を進めることと併せて，対話処理を重要視し，システムが持っている一般的モデルから，使用者の求めている特定のモデルへとグラフを書き換えながら検索を進める手法を論じる．この目的のため，単にマルコフ連鎖グラフの吸収状態だけに注目するのではなく，過渡事象についても関係の重みの強いものを取り出す手法を考察する．

## 2. 事象検索のモデル化

### 2.1 問題の意味

論点を明らかにするため，一つの応用例として FORTRAN プログラムのデバッグ問題を考える．通常，プログラムの実行時の誤りについては，その原因が処理系から直接指摘されないことが多い．例えば“記憶保護領域の侵犯”という誤りの状況によって計算処理が打ち切られているとき，その誤りの源は多数考えられる．配列処理の誤り，引数の結合の誤りによる暴走，等々．さらに，配列処理の誤りについても宣言文が間違っている場合や，添字に現われる変数の値

---

* Interactive Event Searching System on the Basis of Transition Probabilities by Yukuo ISOMOTO (Osaka University Computation Center) and Keiichi YAMAGATA (Department of Precision Engineering, Osaka University).
** 大阪大学大型計算機センター
*** 大阪大学工学部精密工学科

に異常がある場合などがあって，あらゆる原因の可能性を調べあげることは容易ではない.

一般にプログラムの実行時の誤りについては個々のプログラムの構造に依存するため，決定的な誤り追跡の手法がなく，しばしばプログラマの労力が消費される所である．しかしながら経験を積んだプログラマにあっては，その経験に基づいて比較的効率良く誤りの原因を探り出すことができる．このことから，表面に現われる誤りの状況とその原因という因果関係に注目すると，言語処理系の性質とプログラマとしての人間の習性から，確率論的に一般的傾向といったものは把握できるはずである．あくまでも確率論的モデルであるから，一つの誤りの状況に対して複数個の原因が考えられ，また一つの原因に対して表面的に現われるプログラムの振舞いにも多様性がある.

現在，大阪大学大型計算機センターにおいて筆者は上記の確率論的モデルをファイルに登録し，プログラマがこれを手掛りにして能率良く，誤りの原因を探し出せるようなプログラム・デバッグ支援システムを試作した[7]．このシステムは一つ一つのプログラムの構造を解析するわけではないので，使用者とシステムの対話によって，システムからチェック・ポイントとしての事象系列を提示し，使用者が自分のプログラム上での検証結果をシステムに応答するという形式でデバッグが進められる．ここで，システムにとって重要なことは，使用者からの応答を有効に活かして対話の進行に伴って，一般的モデルから個々の使用者に関係する事象だけを抽出し，事後の対話の中に無駄な繰り返しを含まないようにすることである．一回の対話を有効にするため，使用者だけにプログラム検証の負担を負わせるのでなく，システム自身も使用者が現実に抱えている問題により密着した形式へと，モデルを積極的に書き換えて行く機能が主要な論点となる.

### 2.2 定式化

前節で述べたような確率論的重みを持つ事象間の関係(プログラム・デバッグ支援システムでは因果関係)を一般的に取り扱えるようにするため，マルコフ連鎖グラフを導入する．有限事象集合 $V=\{v_1, v_2, \cdots, v_n\}$ を頂点とし，2項関係 $v_iRv_j$ を満たす場合に限り，$v_i$ から $v_j$ へ有向辺をつけ[1]，各有向辺に関係の強さを表す重み $p_{ij}$ をつける．ただし，次式を満たすものとする.

$$\sum_j p_{ij}=1 \qquad (v_i, v_j \in V, v_iRv_j) \tag{2.1}$$

もしも，ある $v_i$ について関係 $R$ を満たす事象 $v_j$ が一つもないときは，便宜上重み1を持つ自己サイクルをつけると，マルコフ連鎖グラフを得る.

つぎに，グラフ理論的考察を行うため，任意の事象 $v_i$ から一段の推移で到達可能な事象集合を $\Gamma(v_i)$ と表し，逆に一段の推移で $v_i$ に到達可能な集合を $\Gamma^{-1}(v_i)$ と表す．これを任意の部分集合 $U\subset V$ についても適用し，$U$ から一段の推移で到達可能な事象集合を $\Gamma(U)$，一段の推移で $U$ へ到達可能な事象集合を $\Gamma^{-1}(U)$ と書く．$k$ 段階の推移については，$\Gamma^k$ を次式で定義する[5].

$$\Gamma^0(v_i)=\{v_i\}, \quad \Gamma^k(v_i)=\Gamma(\Gamma^{k-1}(v_i)), \quad k\geqq 1 \tag{2.2}$$

以上の表現を用いて，

$$\hat{\Gamma}(v_i)=\bigcup_{k=0}^{\infty}\Gamma^k(v_i) \tag{2.3}$$

を定義すれば，$\hat{\Gamma}(v_i)$ は事象 $v_i$ からの到達可能集合を表す．(2.2)式において $\Gamma$ の代りに $\Gamma^{-1}$ を用いて(2.2)，(2.3)と同様に定義すれば，推移の段数を問わないでいつかは $v_i$ に到達可能な事象集合が定義できる．これを $\check{\Gamma}(v_i)$ と書く.

任意の $v_i, v_j$ に対して $v_j\in\hat{\Gamma}(v_i)$ が成り立つとき，かつそのときに限り $v_i\leqq v_j$ であるような関係≦を定義する．この関係は一般的には $V$ 上における擬順序関係である[5]．検索の対象としては半順序関係に注目するものとし，有向辺の向きに従って先行する事象から後続する事象へと探索を進める．グラフ上で複数個の事象が強連結になっている部分は同値類として扱われる[1),5]．この部分を検索の過程から見ると，複数個の事象が房状になっているので，これをクラスターと呼ぶ．クラスターの中には，そこから出て行く有向辺を持つものと，そこから出て行く有向辺を持たない閉じたクラスターとがある．閉じたクラスターと自己サイクルによって単一事象で閉じているものは，後続事象を持たず，従ってこれらはマルコフ連鎖の吸収状態を構成する．このような有向グラフの一例を **Fig. 1** (次頁参照) に示す.

プログラム・デバッグ支援システムの場合，注目している半順序関係は，表面的に現われた誤りの状況とその原因という因果関係である．この場合でも，マルコフ連鎖を考えるときには，多段の因果関係は考慮しないで2つの事象の組（一段だけの関係）に注目してグラフを構成するので，一般的グラフとしてはクラスターが含まれる．しかし，特定の使用者の状況においては因果関係のサイクルはあり得ない．従って，後述する対話処理によって，不要事象を取り除いて行くと

最後にはクラスターは分解される.

さらに端末事象の持つ意味は，それ以上誤りの原因がたどれないこと，すなわちプログラムのバッグの所在を表すことになる.

### 2.3 推移確率行列の役割り

マルコフ連鎖の推移確率行列 $[p_{ij}]$ を構成するに当たり，事象集合を以下のように分類する.

(1) 過渡状態を構成する事象の集合　$T$

(2) 閉じたクラスター　$C_{(1)}, C_{(2)}, \cdots$

(3) 単一で閉じている事象の集合　$U$

ここで閉じたクラスターにつけられた添字は，同値類の番号である. 以後簡単のため，$T$ に属する事象を過渡事象と呼び，また，特に断わらない限り(2)，(3)をひとまとめにして端末事象と呼ぶ. 適当に事象の番号づけを行うことにより，例えば閉じたクラスターが2つある場合，推移確率行列を (2.4) 式の形に整えることができる.

$$P\equiv[p_{ij}]=\left[\begin{array}{c|ccc} Q & \multicolumn{3}{c}{D} \\ \hline 0 & P_1 & 0 & \\ & & P_2 & \\ & 0 & & I \end{array}\right] \tag{2.4}$$

$I$ は単位行列で $U$ における自己サイクルだけの推移を表している. $Q$ は過渡事象相互の間での推移確率を表し，$D$ は過渡事象から端末事象への推移確率を表す. $P_1, P_2$ は閉じたクラスター $C_{(1)}, C_{(2)}$ 内部での推移確率である. 閉じていないクラスターは $T$ の中に含まれているので，それに関する推移確率も $Q$ に含まれており，表面上は区別されない.

いま，使用者によって初期事象 $v_s \in T$ が指定されたとき，使用者の目的に合致するような $v_s$ と端末事象を結ぶ順路を見出すことが，事象検索システムの目的である. 一般にいう情報検索とは内容が異なるので，敢えて区別する意味で事象検索と呼ぶ. 例えば，プログラム・デバッグ支援システムでは，使用者が抱えている表面上の異常状態とその源の原因を結ぶ中間的な誤りの状況が，最終的に求められる順路に対応することになる.

使用者の検索手段の表現法はいろいろ考えられるが，ここでは対話処理を考えているので，使用者自身の判断によるものとする. したがって，検索過程は試行錯誤の繰り返しになるが，システムが事象（または順路）を取り出す規準として推移確率行列が重要な意味を持つ. マルコフ連鎖において端末事象への吸収確率に注目する手法については，すでに論じた[6]. しかし，逆に初期事象の近傍から調べて行きたい場合には，$Q$ の巾乗 $Q^k$ によって $k$ 段階の推移確率を探索の規準とするのが良い. ここで $k$ として小さな値を用いれば，注目している事象に近い事象を調べることになり，$k$ を大きくするにつれて視野が次第に広がって行くことになる.

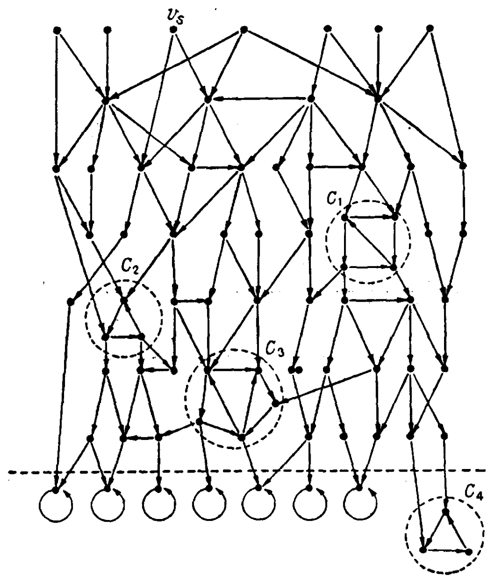


**Fig. 1** A typical directed graph which the event searching system deals with. The probabilistic weights are omitted. The events below a broken line constitute the absorbing states. There are four clusters which are marked with circles.
$C_1$, $C_2$, $C_3$: the clusters having the descendent nodes.
$C_4$: the cluster having no descendent node.

次章以降にまた別の探索規準を導入するが，閉じたクラスターの内部で（即ち一つの同値類の中で），個々の事象を調べるときには，推移確率行列を用いるかまたは定常分布を用いるかの手段しかない[6].

## 3. 対話処理

### 3.1 到達確率の導入

一般に大きなグラフを対象として考えると，端末事象への吸収確率だけに注目した場合，もしもグラフの構造が外向木に近いとすると，多くの端末事象を逐一調べ上げることになる (**Fig. 2**) (次頁参照). また逆に，注目している事象の近傍から調べて行くと，端末

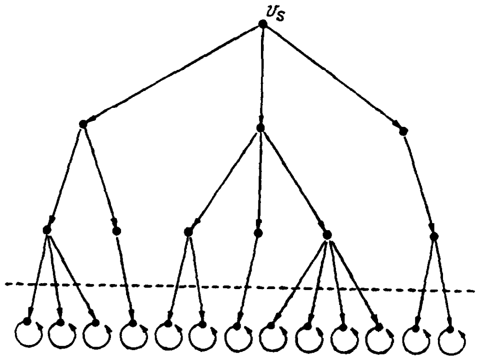


**Fig. 2** A certain special case of the directed graph: out-tree. The events below a broken line constitute the absorbing states.

事象にたどりつく前に多くの対話を繰り返さなければならない．そこで両者の考え方の中間を採るものとして，グラフの全体の中で，現在注目している事象に最も関連の深いものが何であるかを考えることが得策であろう．この目的のために新しく別の探索規準を導入する．

まず過渡事象だけの推移確率行列 $Q$ を用いて，次式を定義する．

$$H\equiv[h_{ij}]\equiv\sum_{k=0}^{\infty}Q^k=[I-Q]^{-1} \tag{3.1}$$

過渡事象だけを考えているので上式は収束する[2]．中にサイクルを含んでいてもよいので，$h_{ij}$ は $v_i$ から出発したとして $v_j$ を訪れる回数の期待値を与える．つぎに，$v_i\in T$ から出発して $k$ 段階推移後に初めて $v_j\in T$ に到達する確率 $g_{ij}^{(k)}$ は次式で与えられる[2]．

$$g_{ij}^{(1)}=q_{ij}^{(1)},$$
$$g_{ij}^{(k)}=q_{ij}^{(k)}-g_{ij}^{(1)}q_{jj}^{(k-1)}-g_{ij}^{(2)}q_{jj}^{(k-2)}-\cdots$$
$$-g_{ij}^{(k-1)}q_{jj}\ (k\geqq 2). \tag{3.2}$$

ここで，$q_{ij}^{(k)}$ は $Q^k$ の第 $i$ 行第 $j$ 列目の要素を表す．(3.1)式が収束するという条件の下で，(3.2)式を $k$ について加え合わせると，次式が得られる．

$$g_{ij}\equiv\sum_{k=1}^{\infty}g_{ij}^{(k)}=\frac{h_{ij}-\delta_{ij}}{h_{jj}},\ (v_i,v_j\in T) \tag{3.3}$$

$g_{ij}$ は推移の段数を問わずに，$v_i$ から出発していつかは $v_j$ に到達する確率を表わす．$\delta_{ij}$ はクロネッカのデルタである．これを $v_i$ からの到達確率と呼ぶことにし，探索の規準としてこの値の大きい $v_j$ を先に取り出すようにする．もしも，グラフが外向木であった場合 (Fig. 2)，推移を重ねるうちに到達確率は減少するので，$g_{ij}$ を規準として隣接事象を順に見て行くことになる．また，**Fig. 3** のような場合，点線の枠内でグラフの構造が複雑になっていても，$v_a, v_b, v_c$ の何れかが比較的早く取り出されることになる．

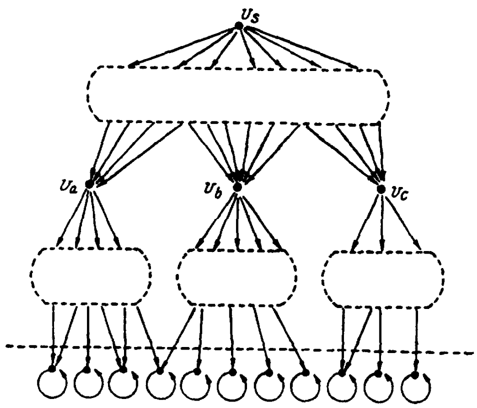


**Fig. 3** A directed graph in which three events $v_a$, $v_b$, $v_c$ occupy the important places.

端末事象への到達確率は吸収確率を用いる．$v_i\in T$ から出発して端末事象 $v_e\in U$ に吸収される確率を $b_{ie}$ とすれば，この値は

$$[b_{ie}]=[I-Q]^{-1}D=HD,\ (v_i\in T, v_e\in U) \tag{3.4}$$

によって得られる．さらに閉じたクラスターを一つの同値類とみなすときは，そのクラスター内の事象への吸収確率の和を用いる．

$$C_{ik}=\sum_{l}b_{il},\ (v_i\in T,\ v_l\in C_{(k)}) \tag{3.5}$$

$C_{(k)}$：閉じたクラスター

### 3.2 使用者の応答

使用者とシステムの対話は，使用者が初期事象 $v_s$ をシステムに投入することから開始される．この段階でシステムは $\hat{\Gamma}(v_s)$ についてだけのグラフを抽出し，隣接行列と推移確率行列を構成する．以後の処理はこの二つの行列に依存して進められるので，主記憶上にグラフそのものを構成する必要はない．

システムはまず $v_s$ から出発するとして，到達確率最大の事象 $v_j$ を (3.3)，(3.4)，(3.5) 式より求める．それぞれ式の形は異なるが，判断の規準のとり方として過渡事象と端末事象の区別はしないので，$\hat{\Gamma}(v_s)$ の抽出時に端末事象が少数個にしぼられているときは，早い時点で端末事象が取り出されることもある．到達確率最大の $v_j$ が見つかると，$v_s$ と $v_j$ とを結ぶ順路の中で，順路に沿う確率の積が最大であるものを求める[6]．ここで得られた順路に沿って，事象系列を使用者に提示する．重要なことは対話によるシステムと使用者の間の情報交換が，事象の系列を介して行われる

ことである．このことは対象のモデルとして半順序関係を考えていることと関連しており，また，一つ一つの事象ごとに対話を繰り返すことを避けて，効率よく処理を進める意味を持つ．

いま $m$ 回目の対話において，新しくシステムから提示された事象系列（実際には事象番号とそれに対応するメッセージの並び）を $S^{(m)}$ とする．

$$S^{(m)} \equiv \{w_1{}^{(m)}, w_2{}^{(m)}, \cdots, w_k{}^{(m)}\} \qquad (3.6)$$
$$m=1, 2, \cdots$$

ここで，$w_k{}^{(m)}$ はシステムによって，$w_1{}^{(m)}$ から出発したときの到達確率が最大であると判断された事象である．これに対して使用者は以下の3種（またはその組み合わせ）の応答が可能である．

ⓐ 事象 $w_a{}^{(m)}$ は現在の検索目的に合致している．

ⓑ 事象 $w_b{}^{(m)}$ は現在の検索目的に合致していない．

ⓒ 提示された $S^{(m)}$ について判断できない．

例えばプログラム・デバッグの支援システムの場合，使用者は $S^{(m)}$ に現われた事象をチェック・ポイントとして，現在抱えている問題のプログラムについて，リスト並びに計算の中間結果を検証し，上記の応答をシステムに返すのである．

実際に誤りの状況が発生していても，それが表面上に出て来ないために判断がつきかねることはしばしば経験する所である．この点を考慮して，ⓒの応答が用意されている．このときシステムは，$w_1{}^{(m)}$ は固定したまま同一のグラフの下で到達確率が $w_k{}^{(m)}$ に対して次に大きい事象を取り出し，改めて順路を決定し使用者に提示する．

システムから提示された $S^{(m)}$ の中に一つでも使用者の目的に合致するものがあれば，ⓐの応答を事象番号を付けてシステムに返す．例えば $w_a{}^{(m)} \in S^{(m)}$ について ⓐの応答をシステムに与えた場合，システムは $w_a{}^{(m)}$ を次の出発点として検索を端末事象に向う方向に進める．即ち $w_1{}^{(m+1)}=w_a{}^{(m)}$ となる (**Fig. 4**)．

ところで，$w_1{}^{(m)}$ から $w_a{}^{(m)}$ に至るすべての事象が目的に合致している時は問題はないが，その途中の事象については合致しない場合がある．つまり，$w_a{}^{(m)}$ は目的に合致していながら，順路については完全な一致が得られていないわけである．この場合，逐一 $w_1{}^{(m)}$ から順路も一致するように検索を進めるか，あるいは順路についてのあいまいさを残したまま，検索を端末事象方向に進めるかは，使用者の指定する事象番号に依存する．換言すれば，探索の仕方について縦

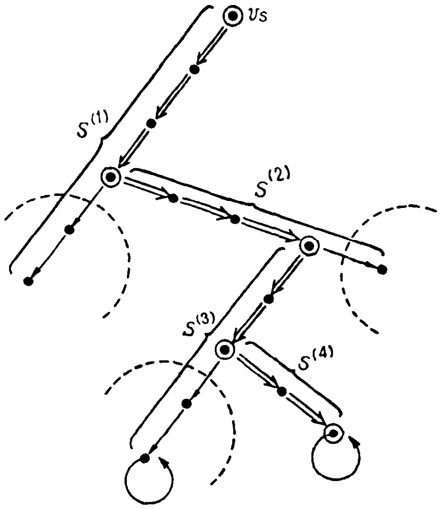


**Fig. 4** Some event sequences picked out through man-machine interactions. Nodes marked ⦿ indicate the user's acknowledge responses.

方向と横方向の自由度があって，使用者はどちらに重きをおくかを選択できるわけである．

つぎに応答ⓑについては，ⓐと組み合わせて用いることができ，提示された事象系列の中で使用者の目的と合わない事象をすべて指定する．複数個の事象を指定して構わない．これによりシステムは，もともとシステム内部に持っている一般的傾向を表すグラフから，現在の使用者の目的に沿うものだけを残し，不要事象を削除してグラフの縮小を図る．当然この処置によって，過渡事象の中のクラスターが分解されることもある．応答ⓐ，ⓑの組み合わせに対してシステムは，不要事象の削除を行った後に，次の検索の出発点をⓐで指定した事象に設定する．

なお，ⓐとⓑで指定する事象は，現在の対話で新しく提示された事象系列 $S^{(m)}$ 以外に，それ以前の対話で提示された事象も対象となり得る．つまり，順路をあいまいにしたままで検索を端末事象に近い所へ進めた後，再び初期事象に近い所へもどって，順路を逐一固めて行く手法が使えるのである．これも使用者の意図によってどこまでもどるかが決められる．もちろん，一回の対話における検索の出発点を，初期事象に近い所へもどしても，不要事象が取り除かれたグラフを対象にすることになるので，様相も変り検索効率もよくなるはずである．

以上のような対話を繰り返すことによって，最終的

に使用者の求める一本の順路が定まる．もちろん，必要な情報が得られた所で対話を打ち切ってもよい．上に述べた機能から明らかなように，たとえ最終的に得られる順路が同一であっても，そこへたどりつく過程には多様性が許されており，使用者のやり易い方法で検索を進めて行くことができる．この点が，この対話システムの大きな特徴になっているのである．

## 4. グラフの書き換え

### 4.1 不要事象の除去

いま，使用者の指定によって取り除くべき事象を $v_d$ とする．このとき，現在扱っているグラフから事象 $v_d$ とそれから出ている有向辺を取り除く．つぎに $v_d$ への有向辺を持つ事象 $v_i \in \Gamma^{-1}(v_d)$ すべてについて $v_d$ への有向辺を取り除く．ここで $v_i$ からの有向辺につけられた確率論的重みの修正が必要になる．もしも，$v_i$ から出ている有向辺が $v_d$ に向うものただ一つであるときは，$v_i$ に重み1の自己サイクルをつけてこれを端末事象とする．この事象は，もともとシステム内部に持っていた一般的モデルでは過渡事象であったものが，現在の使用者の状況に対しては端末事象とみなされる．

$v_i$ から出ている有向辺が複数個あるときは，次式によって推移確率を修正する (**Fig. 5**).

$$p_{ij}'=\frac{p_{ij}}{1-p_{id}} \tag{4.1}$$

$(v_i \in \Gamma^{-1}(v_d),\ v_j \in \Gamma(v_i),\ v_j \neq v_d)$

以上の修正は $\Gamma^{-1}(v_d)$ に属するすべての事象に対して行われる．$v_d$ は過渡事象中のクラスター（出口のあるもの）内の事象であっても，区別して扱う必要はない．無論，複数事象が削除されるときには同一の $v_i$ に対して (4.1) の修正が反復適用されることになる．$\Gamma(v_i)$ の事象がすべて削除されたときは，先の例に準じて $v_i$ を端末事象とする．また，もともと出口のあったクラスターでも，それから出て行く有向辺がすべて取り除かれた場合は，閉じたクラスターとなり，以後は端末事象の扱いとなる．

### 4.2 グラフ全体の縮小

ここで述べることの主眼は，1回の対話におけるグラフの書き換えで，できる限り事象の数を減らし，それ以降でのグラフの再構成操作を簡潔にすることである．前節では，使用者によって明らかに不要であると指定された事象の削除について述べたが，その他にもあってもなくてもよい事象についても取り除くことを

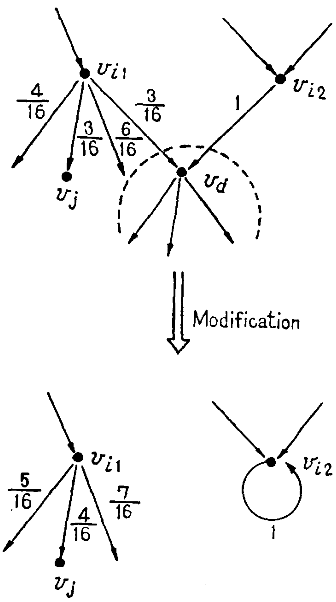


**Fig. 5** An example of the elimination of an event ($v_d$). These fractional numbers denote the probabilistic weights.

考える．

対話の開始時に，使用者によって初期事象 $v_s$ が指定されると，$\hat{\Gamma}(v_s)$ だけが現在の使用者にとって必要である．半順序関係に注目した事象の検索を考えているわけであるから，$v_s$ からの推移の道が存在しない事象は対象からはずして構わない．より具体的に述べれば，グラフの一般的モデルはディスク・ファイルに登録されており，対話の開始時に $\hat{\Gamma}(v_s)$ だけが主記憶装置に持って来られるのである[6]．以後，現在の使用者については，この主記憶上のグラフについて書き換えが実行されて行く．

ある対話において，事象 $v_a$ が使用者の目的に合致していると判断された場合，以後の検索対象としては $\hat{\Gamma}(v_a)$ を考えればよい．ただし，$v_s$ から $v_a$ に至る順路が一義的に決定されていない場合もあるので，この段階での事象集合として以下の $V'$ を考える．

$$V'=\tilde{\Gamma}(v_a)\cup\hat{\Gamma}(v_a) \tag{4.2}$$

$V'$ に属さない事象については，前節で述べた手順によって順に取り除く．関連する重みの修正も行われる．このような事象削除の操作は，ビット・マトリックスで表現された隣接行列を手掛りにして逐次行われ，最後には以後の検索になくてはならない事象だけが残される．修正されたグラフは，単に重みの修正だ

けでなく，過渡事象から端末事象へ，また出口のあるクラスターから出口のない閉じたクラスターへ移行するものもあるので，改めて隣接行列と推移確率行列を作り直す必要がある．

事象集合の大きい間は，グラフの書き換え操作に手間取ることになるが，一回の対話を繰り返すごとに事象の数は大巾に減少して行くので，グラフを積極的に書き換えて行く利点は大きいと考えられる．**3.2** でも述べたように，検索の過程に多くの自由度があるので，初期事象から逐一順路を決めて行くよりも，提示された事象系列について，重要な事象だけに注目して検索を進め，細部の順路については改めて逆もどりをして決定する方が，使用者にとってもシステムにとっても有利であると思われる．

## 5. むすび

本論文で示したマルコフ連鎖グラフの事象検索への応用は，複雑なグラフ上の経路を逐一たどることなしに探索が行われるので，会話処理において大きな効果を発揮する．また，確率論的モデルを扱うことを考慮して，システムと使用者の間で簡潔な応答の組み合わせ方によって，いろいろな検索過程の選択ができることが本システムの特徴になっている．

応用例として述べたプログラム・デバッグ支援システム[7]については，大阪大学大型計算機センターのTSSを用いて実現し，現在試用段階にある．推移確率行列の定め方については，多くの使用者に利用してもらい，統計的データを収集することが必要であろう．

マルコフ連鎖グラフでモデル化できる対象は多いと思われるので，ここに述べた手法はいろいろな応用に使えると思われる．応用の対象が変わる場合，事象対応のメッセージの形式や推移確率行列は当然変化する．しかしながら，事象集合のデータと検索手順の間にマルコフ連鎖という抽象化されたモデルを介在せさることによって，検索手順のプログラムは共通に使うことができる．

今後の課題として，汎用性のあるデータ・ベースに事象集合に関するデータを組み込み，基本データの収集やディスクと主記憶の間のデータ転送の一般化を図り，その土台の上にここに述べた手法を活かすことを検討中である．

なお，本稿をまとめるに当り，査読者から有益な助言を戴いた．ここに記して深謝する次第である．

## 参考文献

1) 小野寛晰：関係の代数，p. 164，教育出版 (1974).
2) 北川敏男編：マルコフ過程，p. 187，情報科学講座，共立出版 (1967).
3) J. L. Baen, D. P. Bovet and G. Estrin: Legality and Other Properties of Graph Models of Computations, J. ACM., Vol. 17, No. 3 (1970).
4) Gerald Salton: Automatic Information Organization and Retrieval, McGraw-Hill Book Company (1968).
5) 尾崎 弘，白山 功：グラフとネットワークの理論，コロナ社 (1973).
Claude Berge: Theory of Graphs and its Applications, Methuen & Co. Ltd, London (1966).
6) 磯本征雄，山県敬一：確率推移検索システム，情報処理，Vol. 17, No. 8, pp. 695～702 (1976).
7) 磯本征雄：質問回答過程とその機械化，第17回プログラミング・シンポジウム報告集，pp. 25～36，情報処理学会 (1976).

(昭和50年3月11日受付)
(昭和52年9月28日再受付)