分散記憶法における探索頻度を考慮した探索路長とその評価†

中 村 良 三†† 松 山 公 一†††

分散記憶法はその衝突の処理の方法によって連鎖法と計算法に大別される．これらの手法における探索路長は見出しの探索頻度が一様であると仮定したときにはすでに求められているが，現実の問題では各見出しの探索頻度は個々に異なるものである．それゆえ，各見出しの探索頻度を考慮した探索路長を求めることができれば，より厳密な探索路長の評価を行うことができる．本論文では，各見出しが探索される確率を考慮に入れた観点から，分離連鎖法における探索路長について議論し，その表現式を導き出す．次に，この表現式で，探索頻度に具体的な確率分布を与えたときの探索路長を示すとともに従来の表現式と比較検討する．

1. はじめに

ブロック構造を許す言語ではブロックは入れ子構造をもっており，外側のブロックで宣言された名前の使用や同じ名前を何度も別々のブロックで宣言して使うことが許されている．

このため，名前の有効範囲規則に従った名前の探索だけでなく，ブロックの開設，閉鎖に伴う名前の格納，削除が混在する[4]．

このような動的状況に対処する名前表の一構成として，分散記憶法のうちの分離連鎖法とスタックを組み合わせて名前表を作成した．この名前表ではスタックにはブロックの開設に伴うブロックの入口の印と名前がどのリストにつながれているかを示す指標を格納し，登録される名前はリストの先頭に順々に挿入する方法を取る．ブロックが閉鎖されたときにはスタックの先頭からブロックの入口の印がくるまで順々に指標を取り出し，その指標が示すリストの先頭の名前を削除する．したがって，この名前表ではブロックが閉鎖されたとき，ブロック単位での名前の削除が容易である．

この名前表で探索路長を評価するとき，各名前の探索頻度が一様であると仮定したときにはすでに論じられているが[1)~3)]，しかし，ブロック構造を許す言語では多重入れ子構造で名前が広域的に使用されるので，探索頻度を一様と仮定して評価することは現実に合わない．したがって，探索頻度を考慮した探索路長を評価できる表現式が必要となる．

本論文では，各見出しが探索される確率を考慮し，その探索路長を評価できる表現式を提案する．次に，すべての見出しが一様に探索されると仮定したとき，従来の表現式とこの提案した表現式から導き出した式と比較し，その相違点を検討する．また，探索頻度に具体的な確率を与えて，その探索路長を数値的に把握する．

2. 従来の表現式

分離連鎖法では n 個の見出しを大きさ m の分散記憶表に一様に分配するとき，分散記憶表にリンクされる任意のリストの長さが k である確率を p_{nk}，この数列 $p_{nk}(k=0,1,\cdots,n)$ の母関数を $P_n(z)$ とすると，これらはそれぞれ次のように表現される[5)]．

$$p_{nk}=\binom{n}{k}\left(\frac{1}{m}\right)^k\left(1-\frac{1}{m}\right)^{n-k} \tag{1}$$

$$P_n(z)=\left(1+\frac{z-1}{m}\right)^n \tag{2}$$

また，表の占有率を $\alpha\left(=\dfrac{n}{m}\right)$，成功探索時の平均探索路長およびその分散を S_n, V_n，同様に不成功探索時のそれらを $\bar{S}_n$，$\bar{V}_n$ とする．

分離連鎖法では見出しの探索頻度が一様と仮定したとき，すでにそれらの探索路長は次のように導き出されている[1)]．

2.1 成功探索の場合

$$S_n=m\sum_k\binom{k+1}{2}p_{nk}/n$$

† The Number of Probes Considering the Frequency Distributions and Its Evaluations by RYOZO NAKAMURA and KIMIKAZU MATSUYAMA (University of Kumamoto).

†† 熊本大学電子計算機室
††† 熊本大学

$$=\frac{m}{n}\left\{p_{n1}+3p_{n2}+\cdots+\frac{n(n+1)}{2}p_{nn}\right\}$$

$$=\frac{m}{n}\left\{\frac{1}{2}P_n''(1)+P_n'(1)\right\}$$

$$=\frac{1}{2}\frac{n-1}{m}+1$$

$$\fallingdotseq\frac{\alpha}{2}+1 \tag{3}$$

$$V_n=\frac{\sum\binom{n}{k_1,\,k_2,\,\cdots,\,k_m}}{m^n}\left\{\frac{\binom{k_1}{2}+\binom{k_2}{2}+\cdots+\binom{k_m}{2}}{n}\right\}^2-\left\{\frac{n-1}{2m}\right\}^2$$

$$=\frac{1}{m^n n^2}\left\{m(m-1)\sum\binom{n}{k_1,\,k_2,\,\cdots,\,k_m}\binom{k_1}{2}\binom{k_2}{2}+m\sum\binom{n}{k_1,\,\cdots,\,k_m}\binom{k_1}{2}^2\right\}-\left\{\frac{n-1}{2m}\right\}^2$$

$$=\frac{1}{m^n n^2}\left\{m(m-1)\left(\frac{1}{4}m^{n-4}n^{\underline{4}}\right)+m(m^{n-4})\left(\frac{1}{4}n^{\underline{4}}+n^{\underline{3}}m+\frac{1}{2}n^{\underline{2}}m^2\right)\right\}-\left\{\frac{n-1}{2m}\right\}^2$$

$$=\frac{(m-1)(n-1)}{2nm^2} \tag{4}$$

ここで $n^{\underline{k}}$ は $\prod_{0\leq j<k}(n-j)$ を表す.

2.2 不成功探索の場合

$$\bar{S}_n=\sum_{k=0}^{n}(k+\delta_k)p_{nk}$$

$$=P_n'(1)+P_n(0)$$

$$=\frac{n}{m}+\left(1-\frac{1}{m}\right)^n$$

$$\fallingdotseq\alpha+e^{-\alpha} \tag{5}$$

$$\bar{V}_n=\sum_{k=0}^{n}(k+\delta_k)^2p_{nk}-\bar{S}_n^2$$

$$=p_{n0}+\sum_{k=1}^{n}k^2p_{nk}-\bar{S}_n^2$$

$$=P_n(0)+P_n'(1)+P_n''(1)-\{P_n'(1)+P_n(0)\}^2$$

$$=\frac{n(m-1)}{m^2}+\left(1-\frac{1}{m}\right)^n\left\{1-\frac{2n}{m}-\left(1-\frac{1}{m}\right)^n\right\}$$

$$\fallingdotseq\alpha+e^{-\alpha}(1-2\alpha-e^{-\alpha}) \tag{6}$$

ただし, $\delta_k=\begin{cases}1 & (k=0)\\ 0 & (k>0)\end{cases}$

3. 提案する表現式

見出しを表に配置する場合，置き場所に自由度がある．したがって，見出しの登録順序によってその配置される場所が異なる．

分離連鎖法では，各見出しの探索頻度が一様である場合には見出しの登録順序に関係なく探索路長が求められる[2),3)].

しかし，見出しの探索頻度を考慮する場合には各見出しの探索頻度および各見出しがリストのどの位置に配置されるか把握する必要がある．

そのため見出しの探索頻度を考慮した探索路長の表現式を導き出すため次のような記号を新たに導入する．

ρ_i: i 番目に登録された見出しが探索される確率．ここに

$$\sum_{i=1}^{n}\rho_i=1 \tag{7}$$

γ_{kj}: 長さ k のリストの先頭から j 番目が探索される確率

登録される各見出しがリストの先頭に順々に挿入されるとすれば，i 番目に登録された見出しが長さ k のリストの先頭から j 番目にリンクされる配置可能回数は $\binom{n-i}{j-1}\binom{i-1}{k-j}$*となる，このときリストの生成可能数は $\binom{n}{k}$ となり，各リストには k 個の見出しがリンクされているので，見出しの重複配置回数は $\binom{n}{k}k/n$，すなわち $\binom{n-1}{k-1}$ となる．

したがって，i 番目に登録された見出しが長さ k のリストの先頭から j 番目に配置される確率は

$$\binom{n-i}{j-1}\binom{i-1}{k-j}\Big/\binom{n-1}{k-1}$$

となる．よって，長さ k のリストの先頭から j 番目が探索される確率 γ_{kj} は次のようになる．

$$\gamma_{kj}=\sum_{i=1}^{n}\binom{n-i}{j-1}\binom{i-1}{k-j}\rho_i\Big/\binom{n-1}{k-1} \tag{8}$$

ただし，$\binom{a}{b}=\begin{cases}0 & (a<b)\\ \text{正} & (a\geq b)\end{cases}$

このとき，長さ k のリストにリンクされているすべて

* 各見出しがリストの後尾に順々に挿入されるときには $\binom{i-1}{j-1}\binom{n-i}{k-j}$ となる．

の見出しが探索される確率は

$$\sum_{j=0}^{k}\gamma_{kj}=1 \tag{9}$$

なる関係が成り立つ．ただし，

$$\gamma_{k0}=\begin{cases}1 & (k=0)\\ 0 & (k>0)\end{cases}$$

すなわち，リストの長さが正の場合 ($k>0$) は

$$\begin{aligned}\sum_{j=1}^{k}\gamma_{kj}&=\sum_{j=1}^{k}\sum_{i=1}^{n}\binom{n-i}{j-1}\binom{i-1}{k-j}\rho_i\Big/\binom{n-1}{k-1}\\
&=\frac{1}{\binom{n-1}{k-1}}\Big\{\rho_1\sum_{j=1}^{k}\binom{n-1}{j-1}\binom{0}{k-j}\\
&\quad+\rho_2\sum_{j=1}^{k}\binom{n-2}{j-1}\binom{1}{k-j}\\
&\quad+\cdots+\rho_n\sum_{j=1}^{k}\binom{0}{j-1}\binom{n-1}{k-j}\Big\}\\
&=\frac{1}{\binom{n-1}{k-1}}\Big\{\rho_1\sum_{j=0}^{k-1}\binom{n-1}{j}\binom{0}{k-1-j}\\
&\quad+\rho_2\sum_{j=0}^{k-1}\binom{n-2}{j}\binom{1}{k-1-j}\\
&\quad+\cdots+\rho_n\sum_{j=0}^{k-1}\binom{0}{j}\binom{n-1}{k-1-j}\Big\}\\
&=\frac{1}{\binom{n-1}{k-1}}\Big\{\rho_1\binom{n-1}{k-1}+\rho_2\binom{n-1}{k-1}+\cdots\\
&\quad+\rho_n\binom{n-1}{k-1}\Big\}\\
&=\rho_1+\rho_2+\cdots+\rho_n\\
&=1\end{aligned}$$

となる．また長さが零のとき，すなわちリストが生成されていないときには $\gamma_{k0}=1$ とする．

さらに，次の記号を導入する．

q_{nk}：任意の長さのリストの先頭から k 番目が探索される確率

この場合は長さが k 以上の各リストの先頭から k 番目が探索される事象と等値であるから先に導入した記号を用いて次のように表現される．

$$q_{nk}=\sum_{i=k}^{n}\gamma_{ik}p_{ni} \tag{10}$$

このとき，

$$\begin{aligned}\sum_{k=0}^{n}q_{nk}&=\sum_{k=0}^{n}\sum_{i=k}^{n}\gamma_{ik}p_{ni}\\
&=\sum_{k=0}^{n}\sum_{j=0}^{k}\gamma_{kj}p_{nk}\\
&=\sum_{k=0}^{n}p_{nk}\\
&=1\end{aligned} \tag{11}$$

が成立する．

ここで，成功探索では探すリストは少なくともひとつ以上の見出しがリンクされているという条件を加味し，探索回数 k を確率変数にとれば，その探索路長の平均，分散は次のように表現される．

$$\begin{aligned}S_n&=\sum_{k=1}^{n}kq_{nk}\Big/\sum_{k=1}^{n}p_{nk}\\
&=\sum_{k=1}^{n}k\sum_{i=k}^{n}\gamma_{ik}p_{ni}\Big/\sum_{k=1}^{n}p_{nk}\end{aligned} \tag{12}$$

$$V_n=\sum_{k=1}^{n}k^2\sum_{i=k}^{n}\gamma_{ik}p_{ni}\Big/\sum_{k=1}^{n}p_{nk}-S_n^2 \tag{13}$$

不成功探索では見出しのひとつもないリストを探すか，リストの最後尾まで探すかのいずれである．前者の探索を1回とすれば，その探索路長の平均，分散は次のようになる．

$$\bar{S}_n=\sum_{k=0}^{n}(k+\delta_k)p_{nk} \tag{14}$$

$$\bar{V}_n=\sum_{k=0}^{n}(k+\delta_k)^2p_{nk}-\bar{S}_n^2 \tag{15}$$

4. 比較検討

ここでは見出しの探索頻度が登録順序に関係なく一様な場合および登録順序に従い半減する場合と調和減少する場合について検討する．

とくに一様な場合は従来の表現式と比較し，その相違点を検討する．

4.1 探索頻度が一様な場合

見出しの探索頻度が一様という仮定であるから，これを提案する表現式で導き出すと次のようになる．探索頻度が一様という仮定と(7)から，$\rho_1=\rho_2=\cdots=\rho_n=\dfrac{1}{n}$ となり，(8)から $\gamma_{kj}=\dfrac{1}{k}$ となる．

よって，成功探索路長 S_n, V_n は次のようになる．

$$\begin{aligned}S_n&=\sum_{k=1}^{n}k\sum_{i=k}^{n}\frac{1}{i}p_{ni}\Big/\sum_{k=1}^{n}p_{nk}\\
&=\frac{\frac{1}{2}\{P_n'(1)+P_n(1)-P_n(0)\}}{\{1-P_n(0)\}}\\
&=\frac{1}{2}\left[\frac{n}{m}\Big/\left\{1-\left(1-\frac{1}{m}\right)^n\right\}+1\right]\\
&\fallingdotseq\frac{1}{2}\{\alpha/(1-e^{-\alpha})+1\}\end{aligned} \tag{16}$$

$$V_n=\sum_{k=1}^{n}k^2\sum_{i=k}^{n}\frac{1}{i}p_{ni}\Big/\sum_{k=1}^{n}p_{nk}-S_n{}^2$$
$$=\frac{1}{\sum_{k=1}^{n}p_{nk}}\left\{p_{n1}+\frac{15}{6}p_{n2}+\cdots+\frac{(n+1)(2n+1)}{6}p_{nn}\right\}-S_n{}^2$$
$$=\frac{1}{6\{1-P_n(0)\}}\{1-P_n(0)+5P_n{}'(1)+2P_n{}''(1)\}$$
$$-\left\{\frac{1}{2\cdot\{1-P_n(0)\}}[P_n{}'(1)+1-P_n(0)]\right\}^2$$
$$=\left\{\frac{n}{m}\frac{n-1}{m}+\frac{n}{m}\right\}\Big/3\left\{1-\left(1-\frac{1}{m}\right)^n\right\}$$
$$-\left(\frac{n}{m}\right)^2\Big/4\left\{1-\left(1-\frac{1}{m}\right)^n\right\}^2-\frac{1}{12}$$
$$\fallingdotseq\alpha(\alpha+1)/3(1-e^{-\alpha})$$
$$-\alpha^2/4(1-e^{-\alpha})^2-\frac{1}{12} \tag{17}$$

不成功探索路長 $\bar{S}_n$, $\bar{V}_n$ は(5), (6)に一致する.

成功探索における従来の表現式(3), (4)といま導出した(16), (17)の相違点について検討する.

初めに, 従来の表現式(3)と(4)が同一の観点から導き出されていることを明確に示すと次のようになる.

$$S_n=\sum_{k_1+\cdots+k_m=n}\left\{\frac{\binom{k_1+1}{2}+\cdots+\binom{k_m+1}{2}}{n}\right\}$$
$$\times\frac{\binom{n}{k_1,\cdots,k_m}^{*,5)}}{m^n}$$
$$=\sum_{k=0}^{n}m\cdot\frac{\binom{k+1}{2}}{n}\frac{\binom{n}{k}(m-1)^{n-k}}{m^n}$$
$$=m\sum_{k=1}^{n}\frac{\binom{k+1}{2}}{n}\cdot\binom{n}{k}\left(\frac{1}{m}\right)^k\left(1-\frac{1}{m}\right)^{n-k}$$
$$=m\sum_{k=1}^{n}\binom{k+1}{2}p_{nk}/n \tag{3'}$$

したがって, (3), (4)の従来の表現式では n 個の見出しを m 個のリストに分配するとき, その任意の分配に対して, n 個のすべての見出しを探索したときの探索路長を見出しの総数で割った, 見出しひとつ当りの探索路長 $\dfrac{\binom{k_1+1}{2}+\cdots+\binom{k_m+1}{2}}{n}$ を確率変数とし, その確率分布を $\dfrac{\binom{n}{k_1,\cdots,k_m}}{m^n}$ としている.

この前提のもとにその平均, 分散の表現式を導出していると考えられる.

しかし, 提案する表現式では探索回数を確率変数と考えているので, 仮にこの前提の形に(3′)を変形するとすれば, (3′)は次のようになる.

$$S_n=m\sum_{k}\binom{k+1}{2}p_{nk}/n$$
$$=\frac{m}{n}\left\{p_{n1}+3p_{n2}+\cdots+\frac{n(n+1)}{2}p_{nn}\right\}$$
$$=\frac{m}{n}\{(p_{n1}+p_{n2}+\cdots+p_{nn})$$
$$+2(p_{n2}+\cdots+p_{nn})+\cdots+np_{nn}\}$$
$$=\sum_{k=1}^{n}k\sum_{i=k}^{n}\frac{1}{\frac{n}{m}}p_{ni} \tag{18}$$

このとき

$$\sum_{k=1}^{n}\sum_{i=k}^{n}\frac{1}{\frac{n}{m}}p_{ni}=1$$

となり, 探索回数 k を確率変数とし, その確率分布は $\sum_{i=k}^{n}\frac{1}{\frac{n}{m}}p_{ni}$ となる.

したがって, k 回の探索を要する見出しは長さ k 以上のリストの先頭から k 番目がいずれも等しい確率 $\frac{1}{\frac{n}{m}}p_{ni}(i\geq k)$ で探索されると解釈される.

この仮定のもとに分散を求めれば次のようになる.

$$V_n=\sum_{k=1}^{n}k^2\sum_{i=k}^{n}\frac{1}{\frac{n}{m}}p_{ni}-S_n{}^2$$
$$=\frac{m}{n}\{p_{n1}+(1+2^2)p_{n2}+\cdots$$
$$+(1+2^2+\cdots+n^2)p_{nn}\}-S_n{}^2$$
$$=\frac{1}{6}\frac{m}{n}\{(p_{n1}+2p_{n2}+\cdots+np_{nn})$$
$$+3(p_{n1}+2^2p_{n2}+\cdots+n^2p_{nn})$$
$$+2(p_{n1}+2^3p_{n2}+\cdots+n^3p_{nn})\}-S_n{}^2$$
$$=\frac{1}{6}\frac{m}{n}\{6P_n{}'(1)+9P_n{}''(1)$$
$$+2P_n{}'''(1)\}-S_n{}^2$$

---

* $\binom{n}{k_1,\cdots,k_m}=\binom{n}{k_1}\binom{n-k_1}{k_2}\binom{n-k_1-k_2}{k_3}\cdots\binom{n-k_1-k_2-\cdots-k_{m-2}}{k_{m-1}}$ なる関係を用いる.

表 1 一様探索時における成功探索路長の比較
Table 1 Comparisons of the number of probes in a successful search for uniform probing.

n	α	(3) $1+\frac{\alpha}{2}$	(4) $\frac{(n-1)(m-1)}{2nm^2}$	(19) $\frac{(n-1)(6m+n-5)}{12m^2}$	(16) $\frac{1}{2}\left(\frac{\alpha}{1-e^{-\alpha}}+1\right)$	(17) $\frac{\alpha(\alpha+1)}{3(1-e^{-\alpha})}-\frac{\alpha^2}{4(1-e^{-\alpha})^2}-\frac{1}{12}$
25	0.5	1.25	0.0094	0.2560	1.135	0.148
50	1.0	1.50	0.0096	0.5635	1.291	0.346
100	2.0	2.00	0.0097	1.3035	1.657	0 892
150	3.0	2.50	0.0097	2.2101	2.079	1.634
200	4.0	3.00	0.0097	3.2835	2.537	2.557
250	5.0	3.50	0.0097	4.5235	3.017	3.649

table size $m=50$

表 2 探索頻度を考慮したときの成功探索路長の比較
Table 2 Comparisons of the number of probes in a successful search under the consideration of probability distribution.

Number of identifiers	Load factor	4.1 uniform		4.2 a) half		4.2 b) harmonic	
		S_n	V_n	S_n	V_n	S_n	V_n
25	0.5	1.13	0.14	1.25	0.26	1.20	0.21
50	1.0	1.29	0.34	1.56	0.63	1.45	0.52
100	2.0	1.65	0.89	2.29	1.54	2.06	1.35
150	3.0	2.07	1.63	3.13	2.59	2.78	2.39
200	4.0	2.53	2.55	4.06	3.68	3.56	3.57
250	5.0	3.01	3.64	5.02	4.75	4.38	4.85

scatter index table size $m=50$

$$=\frac{1}{6}\frac{m}{n}\left\{6\frac{n}{m}+9\frac{n}{m}\frac{n-1}{m}+2\frac{n}{m}\frac{n-1}{m}\frac{n-2}{m}\right\}-\left\{\frac{n-1}{2m}+1\right\}^2$$

$$=\frac{(n-1)(6m+n-5)}{12m^2} \tag{19}$$

結局，従来の表現式(3)，(4)と提案する表現式から導出した(16)，(17)の違いは確率変数の設定の違いにあるものと考えられる．

(4)で導出された分散は $n\gg1, m\gg1$ ならば，$\frac{1}{2m}$ に近似され，mの大きさのみに依存する形となる．一方，提案する表現式から導出された平均成功探索路長およびその分散はいずれも表占有率αの関数として表現することができる．

これらの比較を数値例で**表 1** に示す．

4.2 探索頻度に確率分布を与えた場合

a) 探索頻度が登録順序に従い半減する場合

この仮定では i 番目に登録された見出しが探索される確率は ρ_i であるから，それらは $\rho_1=\frac{c}{1}$，$\rho_2=\frac{c}{2}$，$\cdots$，$\rho_n=\frac{c}{2^{n-1}}$ となる．ただし，$c=\frac{1}{2-2^{1-n}}$ とする．また，長さ i のリストの先頭から k 番目が探索される確率 γ_{ik} は(8)より次のようになる．

$$\gamma_{ik}=\sum_{j=1}^{n}\binom{n-j}{k-1}\binom{j-1}{i-k}\frac{1}{2-2^{1-n}}\frac{1}{2^{j-1}}\Big/\binom{n-1}{i-1}$$

よって，探索路長の平均，分散は次のようになる．

$$S_n=\left\{\frac{1}{2-2^{1-n}}\sum_{k=1}^{n}k\sum_{i=k}^{n}\frac{n}{i}\left(\frac{1}{m}\right)^i\times\left(1-\frac{1}{m}\right)^{n-i}\sum_{j=1}^{n}\binom{n-j}{k-1}\binom{j-1}{i-k}\frac{1}{2^{j-1}}\right\}\Big/\left\{1-\left(1-\frac{1}{m}\right)^n\right\}$$

$$V_n=\left\{\frac{1}{2-2^{1-n}}\sum_{k=1}^{n}k^2\sum_{i=k}^{n}\frac{n}{i}\left(\frac{1}{m}\right)^i\times\left(1-\frac{1}{m}\right)^{n-i}\sum_{j=1}^{n}\binom{n-j}{k-1}\binom{j-1}{i-k}\frac{1}{2^{j-1}}\right\}\Big/\left\{1-\left(1-\frac{1}{m}\right)^n\right\}-S_n^2$$

b) 探索頻度が登録順序に従い調和減少する場合

ここでは $\rho_1=\frac{c}{1}$，$\rho_2=\frac{c}{2}$，$\cdots$，$\rho_n=\frac{c}{n}$ で，$c=\frac{1}{H_n}$ となる．ただし H_n は調和数$\left(H_n=1+\frac{1}{2}+\cdots+\frac{1}{n}\right)$を表す．このとき，$\gamma_{ik}$ は次のようになる．

$$\gamma_{ik}=\sum_{j=1}^{n}\binom{n-j}{k-1}\binom{j-1}{i-k}\frac{1}{j}\frac{1}{H_n}\frac{1}{\binom{n-1}{i-1}}.$$

よって，S_n, V_n は次のように表現される．

$$S_n=\left\{\frac{1}{H_n}\sum_{k=1}^{n}k\sum_{i=k}^{n}\frac{n}{i}\left(\frac{1}{m}\right)^i\left(1-\frac{1}{m}\right)^{n-i}\times\sum_{j=1}^{n}\binom{n-j}{k-1}\binom{j-1}{i-k}\frac{1}{j}\right\}\Bigg/\left\{1-\left(1-\frac{1}{m}\right)^n\right\}$$

$$V_n=\left\{\frac{1}{H_n}\sum_{k=1}^{n}k^2\sum_{i=k}^{n}\frac{n}{i}\left(\frac{1}{m}\right)^i\left(1-\frac{1}{m}\right)^{n-i}\times\sum_{j=1}^{n}\binom{n-j}{k-1}\binom{j-1}{i-k}\frac{1}{j}\right\}\Bigg/\left\{1-\left(1-\frac{1}{m}\right)^n\right\}-S_n^2.$$

ここで，4.1，4.2 節の各場合における数値例を**表 2** に示す．

5. おわりに

本稿では，探索頻度の確率分布を考慮した探索路長を求める一般的な表現式を提案し，2, 3 の確率分布を仮定し，その探索路長を数値的に把握することができた．また，探索頻度が一様である場合の従来の表現式と比較検討し，その前提が大略的であり，その平均は悲観的な，分散は楽観的な表現式であることを示した．

このように探索頻度の確率分布を考慮し，探索路長の分布特性を把握することは意義あるものと考える．

謝辞 確率論の応用研究にご指導，ご討論いただいている城島邦行教授(熊本女子大)に感謝いたします．

また，本稿をまとめるに際し，的確なご指示を下さった査読者の方にお礼を申し上げます．

参 考 文 献

1) Knuth, D. E.: *The Art of Computer Programming*, Vol. 3, *Sorting and Searching*, p. 506, Addison-Wesley, Reading, Mass. (1973).
2) 弓場敏嗣：ハッシングによる見出しの探索技法 I，および II，信学誌，Vol. 63, No. 9, pp. 15-27 (1980).
3) 西原清一：ハッシング技法と応用，情報処理，Vol. 21, No. 9, pp. 980-990 (1980).
4) Gries, D.: *Compiler Costruction for Digital Computer*, pp. 213-230, John Wiley & Son, New York (1971).
5) Feller, W.: *An Introduction to Probability and Its Applications*, Vol. 1, p. 33, John Wiley & Son, New York (1957).

(昭和 57 年 5 月 12 日受付)
(昭和 57 年 9 月 6 日採録)