第3卷 第4期
1983年10月

数 学 研 究 与 评 论
JOURNAL OF MATHEMATICAL
RESEARCH AND EXPOSITION

Vol.3 No.4
Oct. 1983

斐波那契检索算法 F 的重新分析*

刘大有

（吉林大学计算机科学系）

§1 引 言

在多阶段合并（polyphase merge）中，斐波那契数（Fibonacci numbers）起了一个与2的方幂相似的作用．在检索技术方面，斐波那契数的应用起到了“对半检索算法”之替代物的作用．“斐波那契检索算法”是较可取的，它仅仅包含“加”与“减”之运算，连除以2的操作也没有．无疑，斐波那契技术在信息的分类与查找中是很重要的．

在书[1]的第三卷中，对斐波那契检索算法作了描述，并作了分析，在第三卷第416页有用MIX语言编写的程序 F 以及该程序的平均运行时间：

$$(6\phi k/\sqrt{5}-(2+22\phi)/5)u\approx(6.252\log_2 N-4.6)u$$

for a successful search;

$$(6\phi k/\sqrt{5}+(58/(27\phi))/5)u\approx(6.252\log_2 N+5.8)u$$

for a unsuccessful search. (9)

但(9)式的结果是不正确的，本文予以改正．

论文是在吉林大学计算机科学系管纪文教授的指导下完成的，管纪文教授并对全文进行了审阅，作者在此特对管纪文教授表示衷心地感谢．

§2 书[1]第三卷对斐波那契检索算法分析中的错误

1° 不正确的结果．为简便计，仅以成功检索为例．

斐波那契检索算法的平均运行时间的正确结果（§3给出证明）：

$$t_F\approx(7.05\log_2 N-7.21)u$$

对半检索算法 C 的平均运行时间（见书[1]第三卷第413页公式(7)）：

$$t_C\approx(8.5\log_2 N-6)u$$

斐波那契检索算法的平均运行时间的不正确结果（见书[1]第三卷416页公式(9)）：

$$t'_F\approx(6.252\log_2 N-4.6)u$$

2° 书[1]第三卷分析算法 F 时引用的两个错误公式（见书[1]第三卷第669页）：

*1981年12月24日收到．

$$C1_k = C_{k-1};$$

$$C1'_k = C'_{k-1} + F_{k-1}/F_{k+1}$$

3° 以成功检索为例看书[1]第三卷第416页的错误公式(9)是怎样推出的。

先将书[1]第三卷第416页的程序 F (Fibonaccian search) 书写于下:

$$rA \equiv K,\quad rI1 \equiv i,\quad (rI2, rI3) \equiv p-1,\quad (rI3, rI2) \equiv q-1$$

01	START	LDA	K	1					
02		ENT1	F_K	1					
03		ENT2	$F_{K-1}-1$	1					
04		ENT3	$F_{K-2}-1$	1					
05		JMP	F2A	1					
06	F4A	INC1	1,3		18	F4B	INC1	1,2	C2-S-A
07		DEC2	1,3		19		DEC3	1,2	C2-S-A
08		DEC3	1,2		20		DEC2	1,3	C2-S-A
09	F2A	CMPA	KEY,1		21	F2B	CMPA	KEY,1	C
10		JL	F3A		22		JL	F3B	C
11		JE	SUCCESS		23		JE	SUCCESS	C2
12		J2NZ	F4A		24		J3NZ	F4B	C2-S
13		JMP	FAILURE		25		JMP	FAILURE	A
14	F3A	DEC1	1,3		26	F3B	DEC1	1,2	C1
15		DEC2	1,3		27		DEC3	1,2	C1
16		J3NN	F2B		28		J2NN	F2A	C1
17		JMP	FAILURE		29		JMP	FAILURE	1-S-A

对于成功检索有 $S=1$,$A=0$,总运行时间由下式给出:

$$t_F = (8C - 2C1 + 2) \cdot u \tag{I}$$

书[1]卷三在推导中,首先引用了错误公式 $C1_k = C_{k-1}$ (见书[1]卷三第669页习题18之答案),同时又遗漏了公式(I)中的最后一项,即,

$$t_F = (8C - 2C1) \cdot u$$

$$C_k = \frac{3 \cdot k \cdot F_{k+1} + (k-4)F_k}{5(F_{k+1}-1)} - 1 \text{(§3给出此公式之证明)} \tag{II}$$

将 C_k 以(II)式右端的表达式代替之,得

$$t_F = \left(-6 + \frac{24 \cdot k \cdot F_{k+1} + 8 \cdot k \cdot F_k - 32F_k}{5(F_{k+1}-1)} - \frac{6 \cdot k \cdot F_k - 6F_k + 2 \cdot k \cdot F_{k-1} - 10F_{k-1}}{5(F_k - 1)}\right) \cdot u$$

当 k 充分大时,有 $F_k \approx \frac{1}{\sqrt{5}}\phi^k$,且 $F_k \gg 5$,

$$\therefore t_F \approx \left(-6+\frac{24\cdot k+8\cdot k/\phi-32/\phi}{5}-\frac{6\cdot k-6+2\cdot k/\phi-10/\phi}{5}\right)\cdot u$$

注意到$\frac{1}{\phi}=\phi-1$，便有

$$t_F \approx \frac{u\cdot(18k+6k/\phi-24-22/\phi)}{5}=\frac{18k+6k(\phi-1)-24-22(\phi-1)}{5}\cdot u$$

$$=\frac{6k(2+\phi)-(22\phi+2)}{5}\cdot u$$

再注意到$\sqrt{5}\,\phi=\phi+2$，又有

$$t_F \approx (6\phi k/\sqrt{5}-(22\phi+2)/5)\cdot u$$

§3　对斐波那契算法的重新分析

1° 符号含义之说明

C　系指成功检索时在步 $F2$ 所执行的比较次数的平均值。

C'　系指不成功检索时在步 $F2$ 所执行的比较次数的平均值。

$C1$　系指成功检索时，$K<K_i$ 的次数的平均值，即步 $F3$ 执行次数的平均值

$C2\text{-}S$　系指步 $F4$ 执行的平均次数。

A'_k　系指对 k 阶 F 氏树形不成功检索时，达到右失败点的平均次数。

2° 成功检索之分析

显然，$A_k=0$，k系指对 k 阶 Fibonacci tree。

(i) 求 C_k

设$\mathscr{T}_k$表示第 k 阶斐波那契树形中对应所有内节点的比较次数之总和，即 k 阶 F 氏树形的内路径长度加上内节点的总数 $N=F_{k+1}-1$；由 F 氏树形之构造易知$\mathscr{T}_k$满足下面的递推公式：

$$\mathscr{T}_{k+2}=(\mathscr{T}_{k+1}+F_{k+2}-1)+(\mathscr{T}_k+F_{k+1}-1)+1$$

$$\mathscr{T}_{k+2}=\mathscr{T}_{k+1}+\mathscr{T}_k+F_{k+3}-1 \tag{1}$$

$$\mathscr{T}_0=\mathscr{T}_1=0$$

下面将$\mathscr{T}_n$表成 F_{n+1}, F_n, n 之函数(F_{n+1}，F_n 都是斐波那契数)。作发生函数

$$H(z)=\sum_0^{\infty}\mathscr{T}_i z^i \tag{2}$$

$$G(z)=\sum_0^{\infty}F_i z^i \tag{3}$$

由(2)式减去$z\cdot(2)$式，再减去$z^2\cdot(2)$式，

$$z\cdot(2)\text{式：}\ z\cdot H(z)=\mathscr{T}_0 z+\mathscr{T}_1 z^2+\mathscr{T}_2 z^3+\mathscr{T}_3 z^4+\cdots$$

$$z^2\cdot(2)\text{式：}\ z^2\cdot H(z)=\mathscr{T}_0 z^2+\mathscr{T}_1 z^3+\mathscr{T}_2 z^4+\mathscr{T}_3 z^5+\cdots$$

(2)式$-z\cdot$(2)式$-z^2\cdot$(2)式：

$$(1-z-z^2)H(z)=\mathscr{T}_0+(\mathscr{T}_1-\mathscr{T}_0)z+(\mathscr{T}_2-\mathscr{T}_1-\mathscr{T}_0)z^2+(\mathscr{T}_3-\mathscr{T}_2-\mathscr{T}_1)z^3+\cdots$$

$$=0+0+(F_3-1)z^2+(F_4-1)z^3+\cdots$$

$$=(F_3z^2+F_4z^3+F_5z^4+\cdots)-(z^2+z^3+z^4+\cdots)$$

$$H(z)=\frac{z^2}{1-z-z^2}(F_3+F_4z+F_5z^2+\cdots)-\frac{z^2}{1-z-z^2}(1+z+z^2+\cdots)$$

由书[1]第一卷第81页公式（11）知

$$G(z)=z/(1-z-z^2)$$

故又有

$$H(z)=z\cdot G(z)\cdot(F_3+F_4z+F_5z^2+\cdots)-z\cdot G(z)\cdot(1+z+z^2+\cdots) \qquad (4)$$

比较（4）式两边z^n的系数得：

$$\mathscr{T}_n=\Big(\sum_{0\leqslant i\leqslant n+2}F_iF_{n+2-i}\Big)-F_0F_{n+2}-F_1F_{n+1}-F_2F_n-\sum_{0\leqslant j\leqslant n-1}F_j$$

注意到：

$$F_0+F_1+F_2+\cdots+F_{n-2}+F_{n-1}=F_{n+1}-1\ (\text{证明略}),$$

$$F_0=0,\quad F_1=F_2=1,$$

以及书[1]第一卷1.2.8的公式（17）：

$$\sum_{0\leqslant k\leqslant n}F_kF_{n-k}=\frac{n-1}{5}F_n+\frac{2n}{5}F_{n-1}$$

便有，

$$\mathscr{T}_n=\frac{n+1}{5}F_{n+2}+\frac{2n+4}{5}F_{n+1}-F_{n+1}-F_n-F_{n+1}+1$$

$$=\frac{3nF_{n+1}+(n-4)F_n}{5}-(F_{n+1}-1)$$

$$\therefore\quad C_k=\frac{\mathscr{T}_k}{F_{k+1}-1}=\frac{3k\cdot F_{k+1}+(k-4)F_k}{5(F_{k+1}-1)}-1 \qquad (\mathrm{II})$$

（ii）求证 $$(F_{k+1}-1)C1_k=(F_k-1)C_{k-1} \qquad (\mathrm{III})$$

（书[1]第三卷第669页习题18之答案$C1_k=C_{k-1}$是错误的）

用数学归纳法证明。当$k=1$时（Ⅲ）式显然成立，假定$n\leqslant k$时有（Ⅲ）式成立，即

$$\underbrace{(F_{n+1}-1)\cdot C1_n}_{\boxed{\text{n阶F氏树形中$K<K_i$的比较总次数（即F3执行的总次数）}}} = \underbrace{(F_n-1)\cdot C_{n-1}}_{\boxed{\text{$n-1$阶F氏树形比较总次数}}} \quad \text{成立}$$

往证$n=k+1$时，（Ⅲ）式亦真。容易看出有下等式

$$(F_{k+2}-1)C1_{k+1}=(F_{k+1}-1)C1_k+(F_k-1)C1_{k-1}+(F_{k+1}-1)$$

当由k阶F氏树形构造$k+1$阶F氏树形时，$k+1$阶的$k-1$阶右子树形的每一内节点的$K<K_i$的比较次数与$k-1$阶F氏树形的每一内节点的$K<K_i$的比较次数的平均值相同，而$k+1$阶F氏树形的k阶左子树形的每一内节点的$K<K_i$的比较次数的平均值比k阶F氏树形的每一内节点的$K<K_i$的比较次数的平均值均增加1次。

又由归纳假设上式变成：

$$(F_{k+2}-1)\cdot C1_{k+1}=(F_k-1)C_{k-1}+(F_{k-1}-1)C_{k-2}+F_{k+1}-1$$

由$\mathscr{T}_k$之定义立得：

$$(F_{k+2}-1)\cdot C1_{k+1}=\mathscr{T}_{k-1}+\mathscr{T}_{k-2}+F_{k+1}-1$$

由递推公式(1)又得：

$$(F_{k+2}-1)\cdot C1_{k+1}=\mathscr{T}_k=(F_{k+1}-1)C_k,\text{ 证毕。}$$

3° 不成功检索的分析

(i) 与2°同理可证得

$$\left.\begin{array}{l}\mathscr{T}'_{k+2}=\mathscr{T}'_{k+1}+\mathscr{T}'_k+F_{k+3}\\ \mathscr{T}'_0=\mathscr{T}'_1=0\end{array}\right\}$$

$$\mathscr{T}'_k=\frac{3\cdot k\cdot F_{k+1}+(k-4)\cdot F_k}{5}$$

$$C'_k=\frac{\mathscr{T}'_k}{F_{k+1}}=\frac{3\cdot k\cdot F_{k+1}+(k-4)F_k}{5F_{k+1}} \tag{IV}$$

F_{k+1}是外节点总数，

$\mathscr{T}'_k$表示k阶F氏树形的所有外节点的比较总次数。

(ii) 证明

$$A'_k=F_k/F_{k+1} \tag{V}$$

用归纳法证明。当$k=1$时式(V)显然成立。假设当$k\leqslant n$时(V)成立，往证$k=n+1$时式(V)亦真。

事实上，$A'_{n+1}\cdot F_{n+2}=n+1$阶$F$氏树形的右失败点数$=$($n$阶$F$氏树形的右失败点数)$+$($n-1$阶$F$氏树形的右失败点数)$=F_{n+1}\cdot A'_n+F_nA'_{n-1}=F_n+F_{n-1}=F_{n+1}$

$$\therefore A'_{n+1}=\frac{F_{n+1}}{F_{n+2}}$$

(iii) 证明

$$C1'_k\cdot F_{k+1}=C'_{k-1}F_k+F_{k-1} \tag{VI}$$

(书[1]第三卷第669页习题18之答案的公式 $C1'_k=C'_{k-1}+F_{k-1}/F_{k+1}$ 是错误的)

用数学归纳法证明之。因为当$k=0$或$k=1$时，F氏树形简单地是$\boxed{0}$，所以当$k=1$时，有$C1'_1\cdot F_2=0$，$C'_0\cdot F_1+F_0=0$。故有公式(VI)成立。今假定当$k\leqslant n$时，公式(VI)成立，往证$k=n+1$时，公式(VI)亦真。

容易看出有下面的等式，

$$F_{n+2}\cdot C1'_{n+1}=F_{n+1}C1'_n+F_nC1'_{n-1}+F_{n+1}$$

又由归纳假设得：

$$
\begin{aligned}
F_{n+2}C1'_{n+1} &= F_n C'_{n-1} + F_{n-1} + F_{n-1}C'_{n-2} + F_{n-2} + F_{n+1} \\
&= \mathscr{T}'_{n-1} + \mathscr{T}'_{n-2} + F_{n+1} + F_n = \mathscr{T}'_n + F_n \\
&= F_{n+1}C'_n + F_n
\end{aligned}
$$

证毕。

4° 下表给出了书[1]三卷第416页公式（8）的修正。

项 值 统计量	统 计 平 均 值	最 大 值
C	$\phi k/\sqrt{5}+O(1)$	$k-1$
C1	$k/\sqrt{5}+O(1)$	$k-2$
C2-S	$\phi^{-1}k/\sqrt{5}+O(1)$	$\lfloor k/2\rfloor-1$

该表之证明从略。（C，C1，C2-S 是描述成功检索的统计量.）

5° 成功检索的总运行时间.

$$t_F = (8C - 2C1 + 2)\cdot u,$$

$$C1_k = \frac{F_k - 1}{F_{k+1} - 1}C_{k-1}, \qquad C_k = \frac{3\cdot k\cdot F_{k+1} + (k-4)\cdot F_k}{5(F_{k+1}-1)} - 1$$

$$
\begin{aligned}
t_F &= \left(-6 + \frac{24kF_{k+1} + 8kF_k - 32F_k - 6kF_k + 16F_k - 2\cdot k\cdot F_{k-1} + 10\cdot F_{k-1} - 10}{5(F_{k+1}-1)}\right)\cdot u \\
&\approx \left(\frac{22k + 2k\phi - 2k/\phi^2 - 16\phi + 16 + 10/\phi^2}{5} - 6\right)\cdot u \\
&= \frac{18k + 4k\phi + 6 - 26\phi}{5}\cdot u.
\end{aligned}
$$

令 $N = F_{k+1} - 1$，当 k 充分大之时，有 $k \approx \frac{1}{\log_2\phi}\cdot\log_2 N$, $\phi \approx 1.61803$.

∴ $$t_F \approx (7.05\log_2 N - 7.21)\cdot u$$

参 考 文 献

[1] Knuth, D. E. [1973], The Art of Computer Programming, volume 1/Fundamental Algorithms, volume 3/Sorting and Searching, Addison-WesLey, Reading, Mass.