AN APPLICATION OF THE FIBONACCI SEARCH TECHNIQUE TO DETERMINE OPTIMAL. SAMPLE SIZE IN A BAYESIAN DECISION PROBLEM

JEROME D. BRAVERMAN
Rider College
and
DAVID J. TOOF
General Research Corporation

ABOUT THE AUTHORS

Dr. Braverman is Professor and Chairman of the Department of Decisions Sciences and Computers at Rider College, Lawrenceville, New Jersey. He received his Ph.D. from UCLA. He is a charter member of the AIDS and a Senior member of ASQC.

Mr. Toof received his M.S. in Operations Research from Temple University and is a member of the staff of General Research Corporation, McLean, Virginia.

ABSTRACT

In cases where computational difficulties or lack of knowledge about the functional form of a curve preclude the use of analytical methods for determining a maximum, various search techniques can be employed. In Bayesian decision problems, an optimal sample size is based on a maximum expected net gain from sampling. When ENGS is plotted against a range of admissible values of n it is often computationally difficult to determine the maximum. This paper demonstrates how a sequential search technique based on the Fibonacci numbers can be used to determine that value with a minimum number of computations.

AN APPLICATION OF THE FIBONACCI SEARCH TECHNIQUE TO DETERMINE OPTIMAL SAMPLE SIZE IN A BAYESIAN DECISION PROBLEM

One of the most common problems in applied mathematics is the determination of an optimal (max, min) point on the curve of some functional relationship. Sometimes, either because of computational difficulty or lack of knowledge concerning the functional form of the curve itself, it is not feasible to find this optimal point analytically. In such cases, a search technique is a powerful tool. This paper deals with the application of the Fibonacci search technique to the problem of determining the optimal sample size for obtaining additional information in a two-action decision situation with a linear cost function.

THE OPTIMAL SAMPLE SIZE PROBLEM

In any decision problem the question of purchasing additional information is generally approached by comparing the expected value of perfect information, EVPI, with the cost of sampling. EVPI is also equivalent to the cost of uncertainty. Since perfect information can never be obtained from a sample and since it is uneconomical to pay more for information than it could be worth, an amount greater than EVPI should never be spent on sampling. Therefore, the only size samples that would even be considered are those for which the cost of sampling is not greater than EVPI.

For most samples, the cost of a sample of size n can generally be expressed as:

(1) $$C(n) = C_f + nC_v$$

where C_f is the fixed cost of sampling and C_v is the variable cost under the assumption that the incremental cost of each additional sampled unit is the same. The maximum sample size is therefore:

(2) $$n_{\max} \leq \frac{\text{EVPI} - C_f}{C_v} .$$

Any sample size such that $0 \leq n \leq n_{\max}$ is therefore feasible. The problem is to determine the value of n in this range which is optimal. We will designate the optimal value of n as $n*$.

The expected value of the information obtained from any sample of size n can be determined from the expected reduction in the cost of uncertainty that could be achieved with the sample. That is the difference between the EVPI prior to taking the sample and the EVPI after or posterior to the sample. This is computed by means of an extensive form analysis or pre-posterior analysis as described by Sasaki [3], Schlaifer [4], and others. This expected value of sample information is abbreviated EVSI. The expected net gain from sampling, ENGS, is simply EVSI(n) - $C(n)$, that is, the expected value of information from the sample less the cost of obtaining that sample.

The optimal sample size in a decision problem is that value of n, n^*, in the range 0 to n_{max}, for which ENGS is a maximum. Whenever the cost of sampling is high relative to EVPI, n_{max} will be reasonably small and n^* can be determined by simply computing ENGS for every admissible n. However, when the cost of uncertainty is great and sampling costs are not high, this procedure requires a large amount of tedious computations even when performed by computer. Therefore, shortcuts for obtaining n^* are desirable. One such shortcut method would be a Fibonacci search technique.

For the Fibonacci search technique to be effective, it is necessary that ENGS have a single maximum value in the range 0 to n_{max}. Raiffa and Schlaifer [2] have shown that, for two-action problems with linear cost functions, if ENGS has any positive values at all in this range, it will have a single maximum. Consequently, the Fibonacci search technique can be used to find the maximum value of ENGS that corresponds to the optimal sample size in a decision problem of this type.

THE FIBONACCI SEARCH TECHNIQUE

Assume that we are looking for the maximum of a particular curve in the interval (a,b). Then by experimentation we gather information about the curve and reduce the length of our interval of uncertainty. In search techniques, all points of experimentation may be known in advance (preplanned) or information gathered from previous experiments may be used to select the next experimental point (sequential search). The Fibonacci technique is a sequential search technique.

In searching for the maximum of $f(x)$ on the interval (a,b) of length L we perform two experiments as shown in Figure 1. The experiments are performed at points c and d such that the length of $(a,d) = l_1$, and the length of $(c,b) = l_2$. In order to obtain equal intervals, we should let $l_1 = l_2$. If x^* is the true maximum in the interval (a,b), then it follows that:

1. if $f(c) > f(c)$, $x^* \in (a,d)$,
2. if $f(d) > f(c)$, $x^* \in (c,b)$,
3. if $f(c) = f(d)$, $x^* \in (c,d)$.

Case 3 would be extremely rare, and in general either 1 or 2 would occur.

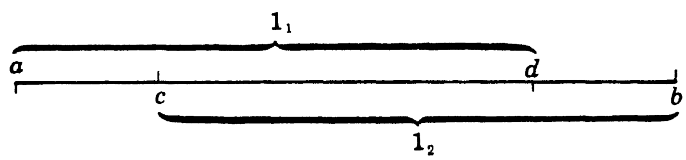

FIGURE 1. Points of Initial Experiments on Interval L

In any case, the new interval of uncertainty would be no greater than l_1. Utilizing the Fibonacci technique after the initial two experiments, it is necessary to perform at most one experiment to determine the next interval.

Now, let us look at the Fibonacci numbers. For any $n \geq 2$, the Fibonacci number, $F_n = F_{n-1} + F_{n-2}$. The Fibonacci series for the first few values of n is:

n:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15 ...
F_n:	1	1	2	3	5	8	13	21	34	55	89	144	233	377	610	987 ...

If we let k be the number of experiments to be performed and L be the length of the initial interval of uncertainty, then after k experiments the interval of uncertainty will be reduced to

$$L_k = \frac{1}{F_k} L.$$

The Fibonacci technique is employed to select sequentially the specific points of experimentation.

AN ILLUSTRATIVE EXAMPLE

As an illustration consider the regret table for a decision problem provided in Table 1 where θ is the unknown decision parameter representing the states of nature and $P_0(\theta)$ is the prior distribution on θ. The EVPI that corresponds to the expected regret of the better act is \$1,050. Assume that $C_f = 0$ and $C_v = \$50$ per unit sampled. From equation (2), n_{max} must be 21 and the initial interval of uncertainty within which n^* must lie is (0,21).

Since we want our search technique to reduce this interval of uncertainty and we know that

$$L_{final} = \frac{1}{F_k} L$$

where k is the number of experimental points that will be taken, we must find the smallest Fibonacci number such that $F_n \geq 21$. From the Fibonacci series above, we can see that $F_7 = 21$ and therefore $k = 7$. This is the maximum number of experiments that it will be necessary to perform.

TABLE 1. Regret Table

$\tilde{\theta}$	$P_0(\tilde{\theta})$	Regret a_1	Regret a_2
0.30	0.70	$1500	$ 0
0.10	0.30	$ 0	$6000
	Expected Regret	$1050	$1800
EVPI = $1050			

The procedure is as follows (the calculation of ENGS for each experiment is provided in the Appendix):

$$L_1 = (0,21), \qquad k = 7;$$

$$L_2 = \frac{F_{k-1}}{F_k} \cdot L_1 = \frac{F_6}{F_7} \cdot L_1 = \frac{13}{21} \cdot 21 = 13.$$

Our initial two points for experimentation will be the two points that are exactly 13 units from the endpoints of the initial interval, L_1. Therefore, L_2 is either (0,13) or (8,21). To determine which of these possible intervals contains n^*, an experiment is conducted at the points $n = 8$ and $n = 13$. That is, ENGS is computed for these two sample sizes resulting in $f(8) = \$34.77$ and $f(13) = -\$50.65$. Since $f(8) > f(13)$, the new interval of uncertainty must be (0,13), that is, $L_2 = (0,13)$.

The next interval,

$$L_3 = \frac{F_5}{F_6} \cdot L_2 = \frac{8}{13} \cdot 13 = 8.$$

By the same procedure employed above, the new interval will be determined by points which are 8 units from tne endpoints of the previous interval. The two new points are $n = 8$ and $n = 5$, and L_3 is either (0,8) or (5,13). It is necessary to compare $f(8)$ with $f(5)$ to determine which of these two possible intervals contains n^*. Since we have already computed $f(8)$, it is now only necessary to determine $f(5)$, which is $97.88, ENGS for a sample of size 5. This property of the Fibonacci search technique which, after the initial two experiments, makes it necessary to conduct only one additional experiment for each additional paired comparison, is one of its great advantages.

Since $f(5) = \$97.88 > f(8) = \$34.77$, the new interval of uncertainty is $L_3 = (0,8)$.

Proceeding,

$$L_4 = \frac{F_4}{F_5} \cdot L_3 = \frac{5}{8} \cdot 8 = 5,$$

and the new interval is either (0,5) or (3,8). Since $f(5) = \$97.88 > f(3) = \$59.22$, the new interval is $L_4 = (3,8)$.

$$L_5 = \frac{F_3}{F_4} \cdot L_4 = \frac{3}{5} \cdot 5 = 3.$$

The new interval is either (3,6) or (5,8).

Since $f(5) = \$97.88 > f(6) = \$80.04$, the new interval is $L_5 = (3,6)$.

$$L_6 = \frac{F_2}{F_3} \cdot L_5 = \frac{2}{3} \cdot 3 = 2.$$

The new interval is either (3,5) or (4,6).

Since $f(5) = \$97.88 > f(4) = \$70.93$, the new interval is (4,6).

$$L_7 = \frac{F_1}{F_2} \cdot L_6 = \frac{1}{2} \cdot 2 = 1.$$

The new interval is one unit (5,5), which is optimal.

At the second-to-last stage, where L was determined to be the interval (4,6), we had already computed $f(4)$, $f(5)$, and $f(6)$, and by comparing the three values could easily see that $n^* = 5$.

Figure 2 shows graphically how the original interval of uncertainty (0,21) was reduced to the interval (4,6). In order to arrive at n^*, ENGS had to be computed for only six values of n, $n = 8$, $n = 13$, $n = 5$, $n = 3$, $n = 6$, and $n = 4$. This is in contrast to having to compute ENGS for 21 integer values in the original interval.

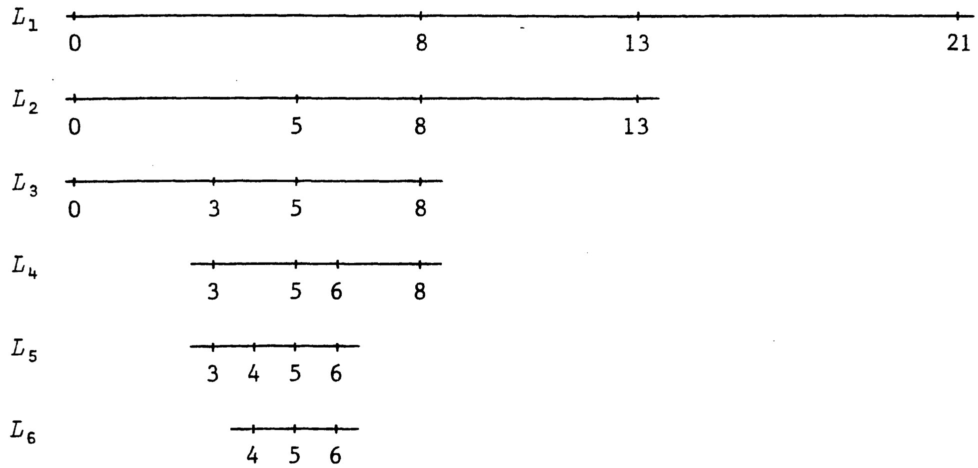

FIGURE 2. Reduction of the Interval of Uncertainty

The computational savings increase even more dramatically as the length of L increases. This is easily seen by looking back to the Fibonacci series and observing, for example, that if $L = (0,987)$, at most 15 experiments would be required, since $F_{15} = 987$.

APPENDIX

Calculation of ENGS

Experiment #1: n = 8

	$\tilde{\theta}$	$P_0(\tilde{\theta})$	$P(x\|\tilde{\theta})$	$P(x \cap \tilde{\theta})$	$P_1(\tilde{\theta})$	Regret a_1	Regret a_2
x = 0:	0.3	0.7	0.06	0.042	0.25	1500	0
	0.1	0.3	0.43	0.129	0.75	0	6000
				0.171		375*	4500
x = 1:	0.3	0.7	0.20	0.140	0.55	1500	0
	0.1	0.3	0.38	0.114	0.45	0	6000
				0.254		825*	2700
x = 2:	0.3	0.7	0.30	0.210	0.82	1500	0
	0.1	0.3	0.15	0.045	0.18	0	6000
				0.255		1230	1080*
x = 3:	0.3	0.7	0.25	0.175	0.95	1500	0
	0.1	0.3	0.03	0.009	0.05	0	6000
				0.184		1426	300*
x = 4:	0.3	0.7	0.14	0.098	0.97	1500	0
	0.1	0.3	0.01	0.003	0.03	0	6000
				0.101		1500	180*
x = 5:	0.3	0.7	0.05	0.035	1	1500	0
	0.1	0.3	0.00	0	0	0	6000
				0.035		1500	0*

(continued)

	$\tilde{\theta}$	$P_0(\tilde{\theta})$	$P(x\|\tilde{\theta})$	$P(x\tilde{\theta})$	$P_1(\tilde{\theta})$	Regret a_1	Regret a_2
$x = 6$:	0.3	0.7	0.01	0.007	1	1500	0
	0.1	0.3	0.00	0	0	0	6000
				0.007		1500	0*
$x = 7$:	0.3	0.7	0.0	0			
	0.1	0.3	0.0	0			
$x = 8$:	0.3	0.7	0.0	0			
	0.1	0.3	0.0	0			

Summary of Posterior Expected Regret for n = 8, X

X	Decision	Marginal Probability	Regret
0	a_1	0.171	375
1	a_1	0.254	825
2	a_2	0.255	1056
3	a_2	0.184	294
4	a_2	0.101	180
5	a_2	0.035	0
6	a_2	0.007	0
7	a_2	0	0
8	a_2	0	0
		Posterior Expected Regret:	615.23

Prior EVPI	1050.00
Post EVPI	-615.23
EVSI (8)	434.77
$C(n = 8)$	-400.00
ENGS (8)	34.77

Experiment #2: n = 13

	$\tilde{\theta}$	$P_0(\tilde{\theta})$	$P(x\|\tilde{\theta})$	$P(x\tilde{\theta})$	$P_1(\tilde{\theta})$	Regret a_1	Regret a_2
$x = 0$:	0.3	0.7	0.10	0.007	0.09	1500	0
	0.1	0.3	0.254	0.076	0.91	0	6000
				0.083		135*	5460
$x = 1$:	0.3	0.7	0.054	0.038	0.26	1500	0
	0.1	0.3	0.367	0.110	0.74	0	6000
				0.148		390*	4440
$x = 2$:	0.3	0.7	0.140	0.098	0.57	1500	0
	0.1	0.3	0.245	0.074	0.43	0	6000
				0.172		855*	2580
$x = 3$:	0.3	0.7	0.22	0.154	0.84	1500	0
	0.1	0.3	0.10	0.030	0.16	0	6000
				0.184		1260	960*
$x = 4$:	0.3	0.7	0.230	0.161	0.85	1500	0
	0.1	0.3	0.028	0.008	0.05	0	6000
				0.169		1425	300*
$x = 5$:	0.3	0.7	0.180	0.126	0.99	1500	0
	0.1	0.3	0.006	0.002	0.01	0	6000
				0.128		1485	60*

(continued)

	$\tilde{\theta}$	$P_0(\tilde{\theta})$	$P(x\|\tilde{\theta})$	$P(x\tilde{\theta})$	$P_1(\tilde{\theta})$	Regret a_1	Regret a_2
x = 6:	0.3	0.7	0.103	0.072	1	1500	0
	0.1	0.3	0	0	0	0	6000
				0.072		1500	0*

For x = 7 through x = 13, $E(R_{a_1}) = 1500E(R_{a_2}) = 0$

Summary of Posterior Expected Regret for n = 13, X

X	Decision	Marginal Probability	Regret
0	a_1	0.083	135
1	a_1	0.148	390
2	a_1	0.172	855
3	a_2	0.184	960
4	a_2	0.169	300
5	a_2	0.128	60
6	a_2	0.072	0
7	a_2	0.031	0
8	a_2	0.0	0
9	a_2	0.0	0
10	a_2	0.0	0
11	a_2	0.0	0
12	a_2	0.0	0
13	a_2	0.0	0
		Posterior Expected Regret:	450.65

Prior EVPI	1050.00
Post EVPI	-450.65
EVSI (13)	599.35
$C(n = 13)$	-650.00
ENGS (13)	- 50.65

Experiment #3: n = 5

	$\tilde{\theta}$	$P_0(\tilde{\theta})$	$P(x\|\tilde{\theta})$	$P(x\tilde{\theta})$	$P_1(\tilde{\theta})$	Regret a_1	Regret a_2
x = 0:	0.3	0.7	0.17	0.119	0.40	1500	0
	0.1	0.3	0.59	0.177	0.60	0	6000
				0.296		600*	3600
x = 1:	0.3	0.7	0.36	0.252	0.72	1500	0
	0.1	0.3	0.33	0.099	0.28	0	6000
				0.351		1080*	1680
x = 2:	0.3	0.7	0.31	0.217	0.91	1500	0
	0.1	0.3	0.07	0.021	0.09	0	6000
				0.238		1365	540*
x = 3:	0.3	0.7	0.13	0.091	0.97	1500	0
	0.1	0.3	0.01	0.003	0.03	0	6000
				0.094		1455	180*
x = 4:	0.3	0.7	0.03	0.04	1	1500	0
	0.1	0.3	0	0	0	0	6000
				0.021		1500	0*

(continued)

	$\tilde{\theta}$	$P_0(\tilde{\theta})$	$P(x\mid\tilde{\theta})$	$P(x\tilde{\theta})$	$P_1(\tilde{\theta})$	Regret a_1	Regret a_2
$x = 5$:	0.3	0.7	0	0		1500	0
	0.1	0.3	0	0		0	6000
							0*

Summary of Posterior Expected Regret for n = 5, X

X	Decision	Marginal Probability	Regret
0	a_1	0.296	600
1	a_1	0.351	1080
2	a_2	0.238	540
3	a_2	0.094	180
4	a_2	0.021	0
5	a_2	0	0
		Posterior Expected Regret:	702.12

Prior EVPI	1050.00
Post EVPI	-702.12
EVSI (5)	347.88
$C(n = 5)$	-250.00
ENGS (5)	97.88

Experiment #4: n = 3

	$\tilde{\theta}$	$P_0(\tilde{\theta})$	$P(x\mid\tilde{\theta})$	$P(x\tilde{\theta})$	$P_1(\tilde{\theta})$	Regret a_1	Regret a_2
$x = 0$:	0.3	0.7	0.34	0.238	0.52	1500	0
	0.1	0.3	0.73	0.219	0.48	0	6000
				0.457		780*	2880
$x = 1$:	0.3	0.7	0.44	0.308	0.81	1500	0
	0.1	0.3	0.24	0.072	0.19	0	6000
				0.380		1215	1140*
$x = 2$:	0.3	0.7	0.19	0.133	0.94	1500	0
	0.1	0.3	0.03	0.009	0.06	0	6000
				0.142		1410	360*
$x = 3$:	0.3	0.7	0.03	0.021	1	1500	0
	0.1	0.3	0	0	0	0	6000
				0.021		1500	0*

Summary of Posterior Expected Regret for n = 3, X

X	Decision	Marginal Probability	Regret
0	a	0.457	780
1	a	0.380	1140
2	a	0.142	360
3	a	0.021	0
		Posterior Expected Regret:	840.78

Prior EVPI	1050.00
Post EVPI	-840.78
EVSI (3)	209.22
$C(n = 3)$	-150.00
ENGS (3)	59.22

Experiment #5: n = 6

	$\tilde{\theta}$	$P_0(\tilde{\theta})$	$P(x\|\tilde{\theta})$	$P(x\tilde{\theta})$	$P_1(\tilde{\theta})$	Regret a_1	Regret a_2
x = 0:	0.3	0.7	0.12	0.084	0.35	1500	0
	0.1	0.3	0.53	0.159	0.65	0	6000
				0.243		525*	3910
x = 1:	0.3	0.7	0.31	0.217	0.67	1500	0
	0.1	0.3	0.35	0.105	0.33	0	6000
				0.322		1005*	1980
x = 2:	0.3	0.7	0.32	0.224	0.88	1500	0
	0.1	0.3	0.10	0.030	0.12	0	6000
				0.254		1320	720*
x = 3:	0.3	0.7	0.19	0.133	0.96	1500	0
	0.1	0.3	0.02	0.006	0.04	0	6000
				0.139		1440	240*
x = 4:	0.3	0.7	0.060	0.0420	0.99	1500	0
	0.1	0.3	0.001	0.0003	0.01	0	6000
				0.0423		1485	60*
x = 5:	0.3	0.7	0.01	0.007	1	1500	0
	0.1	0.3	0	0	0	0	6000
				0.007		1500	0*
x = 6:	0.3	0.7	0	0		1500	0
	0.1	0.3	0	0		0	6000
							0*

Summary of Posterior Expected Regret for n = 6, X

X	Decision	Marginal Probability	Regret
0	a_1	0.243	525
1	a_1	0.322	1005
2	a_2	0.254	720
3	a_2	0.139	240
4	a_2	0.042	60
5	a_2	0.007	0
6	a_2	0	0
		Posterior Expected Regret:	669.96

Prior EVPI	1050.00
Post EVPI	-669.96
EVSI (6)	380.04
$C(n = 6)$	-300.00
ENGS (6)	80.04

Experiment #6: n = 4

	$\tilde{\theta}$	$P_0(\tilde{\theta})$	$P(x\|\tilde{\theta})$	$P(x\tilde{\theta})$	$P_1(\tilde{\theta})$	Regret a_1	Regret a_2
x = 0:	0.3	0.7	0.24	0.168	0.46	1500	0
	0.1	0.3	0.66	0.198	0.54	0	6000
				0.366		690*	3240
x = 1:	0.3	0.7	0.41	0.287	0.77	1500	0
	0.1	0.3	0.29	0.087	0.23	0	6000
				0.374		1155*	1380

(continued)

	$\tilde{\theta}$	$P_0(\tilde{\theta})$	$P(x\|\tilde{\theta})$	$P(x\tilde{\theta})$	$P_1(\tilde{\theta})$	Regret a_1	Regret a_2
$x = 2$:	0.3	0.7	0.26	0.182	0.92	1500	0
	0.1	0.3	0.29	0.015	0.23	0	6000
				0.197		1380	480*
$x = 3$:	0.3	0.7	0.08	0.056	1	1500	0
	0.1	0.3	0.0	0	0	0	6000
				0.056		1500	0*
$x = 4$:	0.3	0.7	0.01	0.007	1	1500	0
	0.1	0.3	0.0	0	0	0	6000
				0.007		1500	0*

Summary of Posterior Expected Regret for n = 4, X

X	Decision	Marginal Probability	Regret
0	a_1	0.366	690
1	a_1	0.374	1155
2	a_2	0.197	480
3	a_2	0.056	0
4	a_2	0.007	0
		Posterior Expected Regret:	779.07

Prior EVPI	1050.00
Post EVPI	-779.07
EVSI (4)	270.93
$C(n = 4)$	-200.00
ENGS (4)	70.93

REFERENCES

1. G. Beveridge and R. Schecter. *Optimization: Theory and Practice.* New York: McGraw-Hill, 1970.
2. H. Raiffa and R. Schlaifer. *Applied Statistical Decision Theory.* Cambridge, Mass.: The MIT Press, 1961.
3. K. Sasaki. *Statistics for Modern Business Decision Making.* Belmont, Calif.: Wadsworth, 1968.
4. R. Schlaifer. *Probability and Statistics for Business Decisions.* New York: McGraw-Hill, 1959.

*****

SIMULTANEOUS TRIBONACCI REPRESENTATIONS

RALPH GELLAR

North Carolina State University, Raleigh, NC 27607

1. INTRODUCTION AND DEFINITIONS

The two-sided sequence $\{t_n\}_{-\infty}^{\infty}$ of Tribonacci numbers is defined by $t_{-1} = 0$, $t_0 = 0$, $t_1 = 1$ and the recursion $t_{n+3} = t_{n+2} + t_{n+1} + t_n$. A Tribonacci representation of the integer a is an expression $a = \Sigma K_i t_i$ where $\{K_n\}_{-\infty}^{\infty}$ is a finitely nonzero sequence of integers.

This paper attempts to generalze to Tribonacci representations some of the results of Robert Silber's and my joint paper [7], "The Ring of Fibonacci Representations." I advise reading that paper before this one because, among other reasons, there one can see how much can be done in the order 2 case.

It is a pleasure to acknowledge here the extensive and essential assistance that Professor Silber gave me in working on the present paper.